

84345

STIC-Biotech/ChemLib

From: Nickol, Gary
Sent: Wednesday, January 15, 2003 12:14 PM
To: STIC-Biotech/ChemLib
Subject: 09/455,486

Please search and interference search the amino acids of SEQ ID NO:6.

Thanks!

Gary Nickol
CM1, AU:1642
Room 8D17, Mailbox 8E12
(703) 305-7143

CRIS

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 1/15/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 18:58:14 ; Search time 41 Seconds
(without alignments)
1475.507 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGMGSPKSLSETCLPN.....ALVLPISIVILDLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgcdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2351	100.0	454	22 AAU10188	Human ORF2 of Six-
2	2351	100.0	454	22 AAE02781	Human six transmem
3	2351	100.0	454	23 ABG61933	Prostate cancer-as
4	2351	100.0	454	23 AAU80190	Human PUMPCn prote
5	2351	100.0	454	23 AAU76538	Tumour-associated
6	2290	97.4	490	22 AAU10187	Human Six-Transmem
7	2036	86.6	419	22 AAU10189	Human ORF3 of Six-
8	1972.5	83.9	576	22 ABG12306	Novel human diagno
9	1972.5	83.9	1273	22 ABG00113	Novel human diagno
10	1738	73.9	450	22 AAE02841	Human STEAP-2 prot

11	1316	56.0	488	22 AAB49483	Rat p-HYDE. Rattu
12	1316	56.0	526	23 ABB83365	Murine Tumour Supp
13	1272	54.1	488	22 AAB85775	Human drug metabol
14	1272	54.1	488	23 ABB83366	Human Tumour Suppr
15	1269	54.0	488	22 AAU10220	Human Six-Transmem
16	1269	54.0	488	22 AAU04564	Human G-protein co
17	1269	54.0	488	22 AAB93224	Human protein sequ
18	1245.5	53.0	487	22 AAB49481	Human p-HYDE. Hom
19	1088	46.3	459	22 AAB74715	Human membrane ass
20	1087	46.2	459	22 AAE02782	Human six transmem
21	1085	46.2	459	22 AAU10190	Human Six-Transmem
22	1082	46.0	458	22 AAE02636	Human STEAP-3 prot
23	1074.5	45.7	456	22 AAB49482	Human p-HYDE 40.
24	901	38.3	173	21 AAY58195	Human STRAP-2 prot
25	844	35.9	179	22 ABB31810	Peptide #4461 enco
26	844	35.9	179	22 ABB37041	Peptide #4547 enco
27	844	35.9	179	22 ABB22356	Protein #4355 enco
28	844	35.9	179	22 AAM57768	Human brain expres
29	844	35.9	179	22 AAM70183	Human bone marrow
30	844	35.9	179	22 AAM18006	Peptide #4440 enco
31	844	35.9	179	22 AAM30517	Peptide #4554 enco
32	844	35.9	179	22 AAM05646	Peptide #4328 enco
33	844	35.9	179	23 ABG39817	Human peptide enco
34	768	32.7	237	22 AAU04565	Human G-protein co
35	736	31.3	141	21 AAY52589	Human prostate gro
36	717	30.5	339	21 AAY58194	Human STRAP-1 prot
37	717	30.5	339	22 AAU69927	Human prostate CDN
38	717	30.5	339	22 AAM78845	Human protein SEQ
39	717	30.5	339	22 AAM01282	P789P amino acid s
40	717	30.5	339	23 ABG61813	Prostate cancer-as
41	717	30.5	339	23 ABB95387	Human P789P protei
42	717	30.5	375	22 AAE02780	Human six transmem
43	704	29.9	374	22 AAM79829	Human protein SEQ
44	695	29.6	339	20 AAW86309	Kidney injury asso
45	435	18.5	129	22 AAB75314	Gene 20 human secr

ALIGNMENTS

RESULT 1

AAU10188

ID AAU10188 standard; Protein; 454 AA.

XX

AC AAU10188;

DT 16-JAN-2002 (first entry)

XX

DE Human ORF2 of Six-Transmembrane Protein of Prostate 1.

XX Human; Six-Transmembrane Protein of Prostate 1; STMP1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF2.

XX Homo sapiens.

PN WO200172962-A2.

XX

PD 04-OCT-2001.

XX

PF 23-MAR-2001; 2001WO-US09410.

XX

PR 24-MAR-2000; 2000US-191929P.

XX

PA (SAAT/) SAATCIOGLU F.

XX

PI Saatioglu F;

XX

DR WPI; 2001-662926/76.

DR N-PSDB; AAS15810, AAS15811.

XX

CC both the sequences differ at several positions.

XX Sequence 454 AA;

SQ Query Match 100.0%; Score 2351; DB 22; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
Db |||||
QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
Db |||||
QY 61 RNPKFASFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
Db |||||
QY 61 RNPKFASFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
Db |||||
QY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAWALQGLPKDASRQVYICSNNIARQQVIE 180
Db |||||
QY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAWALQGLPKDASRQVYICSNNIARQQVIE 180
Db |||||
QY 181 LARQLNFIPIIDLGLSSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
Db |||||
QY 181 LARQLNFIPIIDLGLSSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
Db |||||
QY 241 RNQQSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWPWLETWLQ 300
Db |||||
QY 241 RNQQSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWPWLETWLQ 300
Db |||||
QY 301 CRKQLGLLSFFFAMVHVAYSCLCPMRRSERYLFLNMAYQQVHANIEIENSWNEEEVWRIEMY 360
Db |||||
QY 301 CRKQLGLLSFFFAMVHVAYSCLCPMRRSERYLFLNMAYQQVHANIEIENSWNEEEVWRIEMY 360
Db |||||
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
Db |||||
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
Db |||||
QY 421 EYYRYFTPPNFVLAIVLPSIVILDLLQLCRYPD 454
Db |||||
QY 421 EYYRYFTPPNFVLAIVLPSIVILDLLQLCRYPD 454
Db |||||

RESULT 3
ABG61933

ID ABG61933 standard; Protein; 454 AA.

XX AC ABG61933;

XX XT 15-AUG-2002 (first entry)

DE Prostate cancer-associated protein #134.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

OS Mammalia.

XX WO200230268-A2.

PN 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US32045.

XX 13-OCT-2000; 2000US-0687576.

PR 08-DEC-2000; 2000US-0733288.

PR 08-DEC-2000; 2000US-0733742.

PR 24-JAN-2001; 2001US-263957P.

PR 16-MAR-2001; 2001US-276791P.

PR 16-MAR-2001; 2001US-276888P.

PR 06-APR-2001; 2001US-281922P.

PR 24-APR-2001; 2001US-286214P.

PR 30-APR-2001; 2001US-0847046.

PR 04-MAY-2001; 2001US-288589P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
PI
XX WPI; 2002-471335/50.
DR N-PSDB; ABK92252.
XX

PT Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue -

XX Claim 27; Page 416; 436pp; English.

XX The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridise to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins.

XX Sequence 454 AA;

SQ Query Match 100.0%; Score 2351; DB 23; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
Db |||||
QY 1 MESISMGMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLIRCGYHVIVGS 60
Db |||||
QY 61 RNPKFASFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
Db |||||
QY 61 RNPKFASFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLLVGKILIDVSNM 120
Db |||||
QY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAWALQGLPKDASRQVYICSNNIARQQVIE 180
Db |||||
QY 121 RINQYPESNAEYLASLFPDSLIVKGFNVVSAWALQGLPKDASRQVYICSNNIARQQVIE 180
Db |||||
QY 181 LARQLNFIPIIDLGLSSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
Db |||||
QY 181 LARQLNFIPIIDLGLSSLSSAREIENLPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
Db |||||
QY 241 RNQQSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWPWLETWLQ 300
Db |||||
QY 241 RNQQSDFYKPIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPWPWLETWLQ 300
Db |||||
QY 301 CRKQLGLLSFFFAMVHVAYSCLCPMRRSERYLFLNMAYQQVHANIEIENSWNEEEVWRIEMY 360
Db |||||
QY 301 CRKQLGLLSFFFAMVHVAYSCLCPMRRSERYLFLNMAYQQVHANIEIENSWNEEEVWRIEMY 360
Db |||||
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
Db |||||
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRREFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
Db |||||
QY 421 EYYRYFTPPNFVLAIVLPSIVILDLLQLCRYPD 454
Db |||||
QY 421 EYYRYFTPPNFVLAIVLPSIVILDLLQLCRYPD 454
Db |||||

RESULT 4

AAU80190

ID AAU80190 standard; Protein; 454 AA.

XX

AC AAU80190;
XX 15-JUL-2002 (first entry)
XX Human PUMPCn protein, PRO23203.
DE Human; protein upregulated in metastatic prostate cancer; immunogen;
XX PUMPCn; PRO23203; prostate cancer; cytostatic; gene therapy; transgenic;
KW androgen independent prostate cancer; DNA microarray.
KW Homo sapiens.
XX WO200226822-A2.
OS 04-APR-2002.
XX 26-SEP-2001; 2001WO-US30290.
PN 26-SEP-2000; 2000US-235451P.
XX (GETH) GENENTECH INC.
XX Devaux B, Eberhard D, Goddard A, Godowski PJ, Grimaldi JC;
PI Hillan KJ, Watanabe CK, Wood WI, Yansura DG, Zhang Z;
PI WPI; 2002-383270/41.
DR N-PSDB; ABK50391.
DR
XX New polypeptide termed protein upregulated in metastatic prostate
PT cancer and encoding polynucleotides, useful for identifying polypeptide
PT antagonists for treating prostate cancer .
XX Claim 23; Fig 2; 137pp; English.
PS
XX The invention relates to an isolated human protein upregulated
CC in metastatic prostate cancer (PUMPCn) PRO23203 polypeptide, a
CC sequence 80% identical to PRO23203 and the sequence as encoded by cDNA
CC insert of the vector deposited as ATCC Deposit No. PTA-2513 (DNA185171-
CC 2994) on 26/9/2000. Also included are the polynucleotide encoding
CC the protein (or a DNA sequence 80% identical to the polynucleotide
CC and one that hybridises to complement of the polynucleotide), a vector
CC comprising the polynucleotide, a polynucleotide deposited with ATCC under
CC accession number PTA-2513 (DNA185171-2994), a host cell comprising the
CC vector, preparation of PRO23203, a chimeric molecule comprising PRO23203
CC fused to a heterologous amino acid sequence, an anti-PRO23203 antibody,
CC an agonist/antagonist of PRO23203, and diagnosing the presence of
CC prostate cancer in a mammal by: (a) contacting a microarray diagnostic
CC with a DNA185171-2994 probe, detecting and quantifying hybridisation of
CC DNA185171-2994 probe in prostate cancer tissue compared with normal
CC tissue and determining if DNA185171-2994 is overexpressed; or
CC (b) contacting a tissue of the mammal with an anti-PRO23203 antibody
CC and detecting the binding of the antibody to a component of the tissue,
CC where binding is indicative of the presence of prostate cancer in the
CC mammal. The antibody is useful for treating prostate cancer in mammal
CC which is androgen independent prostate cancer, that has metastasised to
CC cytotoxic agent and the body, where the antibody is not conjugated with a
CC chemotherapeutic agent to the mammal and for diagnosing the presence of
CC prostate cancer in a mammal. The PRO23203 polynucleotide is useful for
CC generating either transgenic animals or knock out animals which in turn
CC are useful in the development and screening of therapeutically useful
CC reagents or for use in gene therapy and for chromosome identification.
CC PRO23203 and polynucleotide may also be used for tissue typing and
CC PRO23203 may also be employed as a therapeutic agents and for screening
CC compounds to identify those that mimic the PRO23203 polypeptides
CC (agonists) or prevent the effect of the PRO23203 polypeptide
CC (antagonists). PRO23203 is further useful for the affinity purification
CC of PRO23203 from recombinant cell culture or natural sources. The
CC present sequence represents PRO23203.
XX Sequence 454 AA;
SQ Query Match 100.0%; Score 2351; DB 23; Length 454;

Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCRGYHVVIGS 60
Db 1 MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCRGYHVVIGS 60
QY 61 RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVCKILIDVSNM 120
Db 61 RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVCKILIDVSNM 120
QY 121 RINQPESNAEYLASLPDPSLIIVKGFNVVSAWALQGPKDASRQVYICSNNIARQQVIE 180
Db 121 RINQPESNAEYLASLPDPSLIIVKGFNVVSAWALQGPKDASRQVYICSNNIARQQVIE 180
QY 181 LARQLNFIPIDGLSSLSSAREIENPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYA 240
Db 181 LARQLNFIPIDGLSSLSSAREIENPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYA 240
QY 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTKYRRFPWPWLETWLQ 300
Db 241 RNOQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTKYRRFPWPWLETWLQ 300
QY 301 CRKQLGLLSFFEFAMVHVAYSCLCPMRRSERVFLNWAYQQVHANIESENWNEEEVWRIEMY 360
Db 301 CRKQLGLLSFFEFAMVHVAYSCLCPMRRSERVFLNWAYQQVHANIESENWNEEEVWRIEMY 360
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIIYGWKRAFE 420
Db 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRBESFIQSTLGYVALLISTFHVLIIYGWKRAFE 420
QY 421 EYYRYFTPPNFVLALVLPISIVILDLLQLCRYPD 454
Db 421 EYYRYFTPPNFVLALVLPISIVILDLLQLCRYPD 454
RESULT 5
AAU76538
ID AAU76538 standard; Protein; 454 AA.
XX AAU76538;
AC AAU76538;
XX 05-JUN-2002 (first entry)
DT Tumour-associated antigenic target protein, TAT138.
XX Tumour-associated antigenic target protein, TAT138.
KW TAT138; Tumour-associated Antigenic Target; tumour;
KW breast cancer; colorectal cancer; lung cancer; ovarian cancer;
KW central nervous system cancer; liver cancer; bladder cancer; melanoma;
KW pancreatic cancer; leukaemia; gene therapy.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Modified-site 8..13
FT /note= "N-myristoylation site"
FT Modified-site 24..29
FT /note= "N-myristoylation site"
FT Region 29..32
FT /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT Modified-site 34..39
FT /note= "N-myristoylation site"
FT Modified-site 193..198
FT /note= "N-myristoylation site"
FT Domain 210..230
FT /note= "Transmembrane domain"
FT Domain 257..277
FT /note= "Transmembrane domain"
FT Modified-site 256..259
FT /note= "Asn is N-glycosylated"
FT Modified-site 274..279
FT /note= "N-myristoylation site"

FT Domain 299..319 /note= "Transmembrane domain"
FT Domain 259..379 /note= "Transmembrane domain"
FT Domain 393..413 /note= "Transmembrane domain"
FT Modified-site 416..424 /note= "Tyrosine kinase phosphorylation site"
FT Domain 428..448 /note= "Transmembrane domain"
FT

XX WO200216429-A2.
PN

XX 28-FEB-2002.
PD

XX 22-JUN-2001; 2001WO-US20118.
PF

XX 24-AUG-2000; 2000WO-US23328.
PR

PR 26-SEP-2000; 2000US-235451P.
PR

PR 01-DEC-2000; 2000WO-US32678.
PR

PR 28-FEB-2001; 2001WO-US06520.
PR

PR 01-MAR-2001; 2001WO-US06666.
PR

XX (GETH) GENENTECH INC.
PA

XX Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;
PI

PI Wood WI, Wu TD, Zhang Z;
PI

XX WPI; 2002-280917/32.
DR

DR N-PSDB; ABK11093.
DR

XX Novel isolated tumour-associated antigenic target polypeptides which are
PT useful as targets for cancer therapy and diagnosis in mammals -
PT

XX Claim 12; Fig 10; 121pp; English.
PS

XX The invention relates to an isolated tumour-associated antigenic target
CC polypeptide (TAT) (I), specifically TAT134-TAT138 polypeptides, and the
CC polynucleotides (II) encoding them. (II) is useful for diagnosing the
CC presence of a tumour in a mammal, where the level of expression of (II)
CC is indicative on the presence of tumour in the mammal from which the
CC test sample was obtained. Antibody to (I) is useful for killing a cancer
CC cell (e.g. breast cancer cell, a colorectal cancer cell, a lung cancer
CC cell, an ovarian cancer cell, a central nervous system (CNS) cancer cell,
CC a liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a
CC melanoma cell or a leukaemia cell) that expresses (I). Oligonucleotides
CC hybridising to (II) are useful as diagnostic probes, antisense
CC oligonucleotide probes or for encoding fragments of full length TAT
CC polypeptide. (II) is also useful in chromosome and gene mapping and in
CC the generation of antisense RNA and DNA probes, for constructing
CC hybridisation probes for mapping the gene encoding TAT and for genetic
CC analysis of individuals with genetic disorders. (II) is also useful for
CC generating either transgenic animals or knockout animals, and in gene
CC therapy. The TAT polypeptides and nucleic acids may also be used for
CC tissue typing and the TAT polypeptides are useful for screening
CC compounds that mimic the TAT polypeptide (agonist) or prevent the
CC effect of TAT polypeptide (antagonist). The antibody is useful for
CC staging TAT polypeptide-expressing cancers, purifying or
CC immunoprecipitating TAT polypeptide from cells, for detection and
CC quantitation of TAT polypeptide in vitro, e.g., in an enzyme linked
CC immunosorbent assay (ELISA) or Western blot. The antibodies are also
CC useful for treating a TAT-expressing cancer or alleviating one or more
CC symptoms of cancer in a mammal. The present sequence represents the
CC amino acid sequence of TAT138.
XX

SQ Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 23; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMMGSPKSLSETCLPNGINGIKDARKVTVGVIGSGDFAKSLTIRLCGYHVVIGS 60
|||||

Db 1 MESISMMGSPKSLSETCLPNGINGIKDARKVTVGVIGSGDFAKSLTIRLCGYHVVIGS 6Q
QY 61 RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGILIDVSNM 120
|||||
Db 61 RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGILIDVSNM 120
|||||
QY 121 RINQYPPESNAEYLASLFPDSLIVKGFNVVSAWALQLPKPDASRQVYICSNNIQARQOVIE 180
|||||
Db 121 RINQYPPESNAEYLASLFPDSLIVKGFNVVSAWALQLPKPDASRQVYICSNNIQARQOVIE 180
|||||
QY 181 LARQLNFIPIIDLGLSSLSSAREIENPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
|||||
Db 181 LARQLNFIPIIDLGLSSLSSAREIENPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA 240
|||||
QY 241 RNQQSDFYKPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYRRFPWLETWLQ 300
|||||
Db 241 RNQQSDFYKPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYRRFPWLETWLQ 300
|||||
QY 301 CRKQLGLSFFPFAMVHVAYSCLCLPMRRSERYLFLNMAYQQVHANIEENSWNEEEVWRIEMY 360
|||||
Db 301 CRKQLGLSFFPFAMVHVAYSCLCLPMRRSERYLFLNMAYQQVHANIEENSWNEEEVWRIEMY 360
|||||
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRFFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
|||||
Db 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRFFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
|||||
QY 421 EYYRFRYTPPNFVLALVLPISIVILDLLQLCRYPD 454
|||||
Db 421 EYYRFRYTPPNFVLALVLPISIVILDLLQLCRYPD 454
|||||

RESULT 6

AAU10187

ID AAU10187 standard; Protein; 490 AA.

XX

AC AAU10187;

XX

DT 16-JAN-2002 (first entry)

XX

DE Human Six-Transmembrane Protein of Prostate 1, STMPL.

XX

KW Human; Six-Transmembrane Protein of Prostate 1; STM;1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Domain 209..230

FT /label= Transmembrane_domain_1

FT Domain 255..273

FT /label= Transmembrane_domain_2

FT Domain 304..325

FT /label= Transmembrane_domain_3

FT Domain 360..380

FT /label= Transmembrane_domain_4

FT Domain 393..415

FT /label= Transmembrane_domain_5

FT Domain 432..452

FT /label= Transmembrane_domain_6

XX

PN WO200172962-A2.

XX

PD 04-OCT-2001.

XX

PF 23-MAR-2001; 2001WO-US09410.

XX

PR 24-MAR-2000; 2000US-191929P.

XX

PA (SAAT/) SAATCIOGLU F.

PI Saatcioglu F;
XX
DR WPI; 2001-662926/76.
DR N-PSDB; AAS15793, AAS15801, AAS15802.
XX
PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids -
XX
PS Claim 1; Fig 4B; 114pp; English.
XX
CC The invention relates to substantially pure prostate-specific or
CC testis-specific polypeptides and the nucleic acids encoding them.
CC Also included are vectors and host cells expressing the proteins, a
CC transgenic animal expressing the protein, antibodies against the
CC proteins, probes for detecting the nucleic acids, antisense molecules
CC for the nucleic acids and methods of isolating modulators of the
CC proteins. Compounds that modulate the prostate specific or testis
CC specific polypeptide are useful to diagnose, prevent or treat disorders
CC of the testis or prostate particularly prostate cancer, benign
CC prostatic hyperplasia, acute prostatitis, testicular cancer,
CC cryptorchidism, undescended, retractile, ascending or vanished
CC testis. Other proliferative disorders for which the modulators may be
CC used include lymphoma, leukaemia, melanoma, ovarian cancer, breast
CC cancer, pancreatic cancer, liver cancer and lung cancer. The
CC present sequence represents a prostate specific protein, Six-
CC Transmembrane Protein of Prostate 1, STM;1.
XX
SQ Sequence 490 AA;
Query Match 97.4%; Score 2290; DB 22; Length 490;
Best Local Similarity 99.8%; Pred. No. 5.5e-238;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
Db |
QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
Db |
QY 61 RNPKFASFEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNM 120
Db |
QY 61 RNPKFASFEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNM 120
Db |
QY 121 RINQYPESNAEYLASLFPDLSLVKGFNVVSAWALQGPDKASRQVYICSNNIARQQVIE 180
Db |
QY 121 RINQYPESNAEYLASLFPDLSLVKGFNVVSAWALQGPDKASRQVYICSNNIARQQVIE 180
Db |
QY 181 LARQLNFIPIDLGLSSLSSAREIENPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db |
QY 181 LARQLNFIPIDLGLSSLSSAREIENPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db |
QY 241 RNQQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPFWLETWLQ 300
Db |
QY 241 RNQQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPFWLETWLQ 300
Db |
QY 301 CRKQLGLLSFFFAMVHVAYSCLPMRRSERYLFNLMAYQQVHANIEWSNNEEVWRIEMY 360
Db |
QY 301 CRKQLGLLSFFFAMVHVAYSCLPMRRSERYLFNLMAYQQVHANIEWSNNEEVWRIEMY 360
Db |
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRREFSFQSTLGYVALLISTFHVLIYGWKRAFE 420
Db |
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWRREFSFQSTLGYVALLISTFHVLIYGWKRAFE 420
Db |
QY 421 EYYRYFTPPNFVLALVLPISIVIL 444
Db |
QY 421 EYYRYFTPPNFVLALVLPISIVIL 444
Db |
RESULT 7
AAU10189
ID AAU10189 standard; Protein; 419 AA.
XX
AC AAU10189;
XX

DT 16-JAN-2002 (first entry)
XX
DE Human ORF3 of Six-Transmembrane Protein of Prostate 1.
XX
KW Human; Six-Transmembrane Protein of Prostate 1; STM;1; prostate cancer;
KW benign prostatic hyperplasia; acute prostatitis; testicular cancer;
KW cryptorchidism; testicular disorder; proliferative disorder; lymphoma;
KW leukaemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;
KW liver cancer; lung cancer; cytostatic; ORF3.
XX
OS Homo sapiens.
XX
PN WO200172962-A2.
XX
PD 04-OCT-2001.
XX
PF 23-MAR-2001; 2001WO-US09410.
XX
PR 24-MAR-2000; 2000US-191929P.
XX
PA (SAAT/) SAATCIOGLU F.
XX
PI Saatcioglu F;
XX
XX WPI; 2001-662926/76.
DR N-PSDB; AAS15813.
XX
PT New polynucleotide for the diagnosis, prevention and treatment for
PT prostate and testis disorders, particularly prostate cancer, comprises
PT prostate-specific or testis-specific nucleic acids -
XX
PS Claim 1; Fig 4K; 114pp; English.
XX
CC The invention relates to substantially pure prostate-specific or
CC testis-specific polypeptides and the nucleic acids encoding them.
CC Also included are vectors and host cells expressing the proteins, a
CC transgenic animal expressing the protein, antibodies against the
CC proteins, probes for detecting the nucleic acids, antisense molecules
CC for the nucleic acids and methods of isolating modulators of the
CC proteins. Compounds that modulate the prostate specific or testis
CC specific polypeptide are useful to diagnose, prevent or treat disorders
CC of the testis or prostate particularly prostate cancer, benign
CC prostatic hyperplasia, acute prostatitis, testicular cancer,
CC cryptorchidism, undescended, retractile, ascending or vanished
CC testis. Other proliferative disorders for which the modulators may be
CC used include lymphoma, leukaemia, melanoma, ovarian cancer, breast
CC cancer, pancreatic cancer, liver cancer and lung cancer. The
CC present sequence is prostate specific protein, Six-Transmembrane
CC Protein of Prostate 1, STM;1, ORF3.
XX
SQ Sequence 419 AA;
Query Match 86.6%; Score 2036; DB 22; Length 419;
Best Local Similarity 99.7%; Pred. No. 1.1e-210;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
Db |
QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS 60
Db |
QY 61 RNPKFASFEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNM 120
Db |
QY 61 RNPKFASFEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNM 120
Db |
QY 121 RINQYPESNAEYLASLFPDLSLVKGFNVVSAWALQGPDKASRQVYICSNNIARQQVIE 180
Db |
QY 121 RINQYPESNAEYLASLFPDLSLVKGFNVVSAWALQGPDKASRQVYICSNNIARQQVIE 180
Db |
QY 181 LARQLNFIPIDLGLSSLSSAREIENPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db |
QY 181 LARQLNFIPIDLGLSSLSSAREIENPLRLFTLWRGPVVAISLATEFFLYSFVRDVIHPYA 240
Db |
QY 241 RNQQSDFYKIPIEIVNKTLPIVAITLLSLVYLAGLLAAAYQLYYGTYKRRFPFWLETWLQ 300

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 18:58:14 ; Search time 18 Seconds
(without alignments)
742.111 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGSPKSLSETCLPN.....ALVLPISIVILDLLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	901	38.3	173	4 US-09-323-873A-8	Sequence 8, Appli
2	736	31.3	141	3 US-09-083-521-1	Sequence 1, Appli
3	717	30.5	339	4 US-09-323-873A-2	Sequence 2, Appli
4	144.5	6.1	227	4 US-09-655-270A-15	Sequence 15, Appli
5	144.5	6.1	227	4 US-09-651-941-17	Sequence 17, Appli
6	144.5	6.1	227	4 US-09-955-597-17	Sequence 17, Appli
7	107.5	4.6	695	1 US-08-487-886-2	Sequence 2, Appli
8	107.5	4.6	695	3 US-08-482-855-2	Sequence 2, Appli
9	107.5	4.6	695	4 US-08-474-986-2	GENERAL INFORMA
10	107	4.6	34	4 US-09-323-873A-20	Sequence 20, Appli
11	101	4.3	476	4 US-09-316-083-3	Sequence 3, Appli
12	99	4.2	531	2 US-08-724-974A-2	Sequence 2, Appli
13	97.5	4.1	390	3 US-08-460-576-2	Sequence 2, Appli
14	97	4.1	692	4 US-07-757-342D-6	Sequence 6, Appli
15	93	4.0	365	2 US-08-724-974A-3	Sequence 3, Appli
16	91.5	3.9	495	4 US-08-985-343-2	Sequence 2, Appli
17	91	3.9	808	4 US-09-134-001C-3105	Sequence 3105, Ap
18	89.5	3.8	495	1 US-07-841-997A-2	Sequence 2, Appli
19	89.5	3.8	495	1 US-08-290-301-2	Sequence 2, Appli
20	89.5	3.8	495	2 US-08-588-983-2	Sequence 2, Appli
21	89.5	3.8	495	2 US-08-588-976-2	Sequence 2, Appli
22	89.5	3.8	495	4 US-09-013-598-2	Sequence 2, Appli
23	89.5	3.8	801	4 US-09-134-001C-5584	Sequence 5584, Ap
24	89.5	3.8	968	3 US-08-651-999A-7	Sequence 7, Appli
25	89.5	3.8	968	4 US-09-385-752-7	Sequence 7, Appli
26	89	3.8	1582	2 US-08-404-531B-9	Sequence 9, Appli
27	89	3.8	1582	3 US-08-476-900A-9	Sequence 9, Appli

28	89	3.8	1582	3 US-08-488-546A-9	Sequence 9, Appli
29	89	3.8	1582	4 US-08-726-320-5	Sequence 5, Appli
30	89	3.8	1582	4 US-09-208-716-5	Sequence 5, Appli
31	88.5	3.8	274	4 US-09-134-001C-3924	Sequence 3924, Ap
32	88.5	3.8	591	1 US-08-484-840-2	Sequence 2, Appli
33	88.5	3.8	591	1 US-08-483-094-2	Sequence 2, Appli
34	87	3.7	467	2 US-08-805-118-3	Sequence 3, Appli
35	87	3.7	467	4 US-09-391-958-3	Sequence 3, Appli
36	86.5	3.7	699	1 US-08-348-006B-7	Sequence 7, Appli
37	86.5	3.7	699	2 US-08-800-825A-7	Sequence 7, Appli
38	86.5	3.7	699	4 US-09-158-657-7	Sequence 7, Appli
39	85.5	3.6	861	3 US-09-022-875-2	Sequence 2, Appli
40	85	3.6	682	4 US-09-134-001C-5250	Sequence 5250, Ap
41	85	3.6	1580	4 US-08-726-320-1	Sequence 1, Appli
42	85	3.6	1580	4 US-09-208-716-1	Sequence 1, Appli
43	85	3.6	1581	2 US-08-404-531B-6	Sequence 6, Appli
44	85	3.6	1581	3 US-08-476-900A-6	Sequence 6, Appli
45	85	3.6	1581	3 US-08-488-546A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-323-873A-8
; Sequence 8, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Safran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-8

Query Match 38.3%; Score 901; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 5.8e-87;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	246	DFYKIP	IEIVNKTLP	IVAITLLSLVYLAGLLAAAYQLYYGTYKRRFP	PWLETWLQCRKQL	305
Db	1	DFYKIP	IEIVNKTLP	IVAITLLSLVYLAGLLAAAYQLYYGTYKRRFP	PWLETWLQCRKQL	60
QY	306	GLLSFFF	AMVHVAYS	CLCPNRRSERYLFLNMAYQQVHAN	ENSWNEEEVWRIEYISFGI	365
Db	61	GLLSFFF	AMVHVAYS	CLCPNRRSERYLFLNMAYQQVHAN	ENSWNEEEVWRIEYISFGI	120
QY	366	MSLGLLS	LLAVTSIP	SVSNALNWR	REFSFIQSTLGYVALLISTFHVLIYGWKRA	418
Db	121	MSLGLLS	LLAVTSIP	SVSNALNWR	REFSFIQSTLGYVALLISTFHVLIYGWKRA	173

RESULT 2
US-09-083-521-1
; Sequence 1, Application US/09083521
; Patent No. 6048970
; GENERAL INFORMATION:

APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,521
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0527 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSTUT10
CLONE: 1691243
US-09-083-521-1

Query Match 31.3%; Score 736; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 1e-69;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 314 MVHVAYSLCLPMRSEYFLNMQVQVHANIEVWRIEMVIFSGIMSLGLLSL 373
Db 1 MVHVAYSLCLPMRSEYFLNMQVQVHANIEVWRIEMVIFSGIMSLGLLSL 60
QY 374 LAVTSIPSVSNALNWREFSFIQSTGLGVALLISTFHVLIYGWKRAFEFEEYRFTPPNFV 433
Db 61 LAVTSIPSVSNALNWREFSFIQSTGLGVALLISTFHVLIYGWKRAFEFEEYRFTPPNFV 120
QY 434 LALVLPISVILDLLQLCRYPD 454
Db 121 LALVLPISVILDLLQLCRYPD 141

RESULT 3
US-09-323-873A-2
Sequence 2, Application US/09323873A
Patent No. 6329503
GENERAL INFORMATION:
APPLICANT: Daniel E. Afar
APPLICANT: Rene S. Hubert
APPLICANT: Kahan Leong
APPLICANT: Arthur B. Raitano
APPLICANT: Douglas C. Saffran
APPLICANT: Steve Chappell Mitchell
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
FILE REFERENCE: 129.16US02
CURRENT APPLICATION NUMBER: US/09/323,873A
CURRENT FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: 60/087,520

PRIOR FILING DATE: 1998-06-01
PRIOR APPLICATION NUMBER: 60/091,183
PRIOR FILING DATE: 1998-06-30
NUMBER OF SEQ ID NOS: 32
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 339
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-323-873A-2
Query Match 30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 3.9e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;
QY 208 LFTLWRGPVVVAISLATFFFLYSFVRDVIHPYARNQQSDFYKIPVIAITLL 267
Db 67 LFPQWHLPIKIAAIIASLTFLYTLREVIHPLATSHQQYFYKIPILVINKVLPWVSITLL 126
QY 268 SLVYLAGLLAAAYQLYYGTKYRPPWLETLQCRKQLGLLSFFFAMVHVAYSCLCLPMRR 327
Db 127 ALVYLPGVIAIVQLHNGTKYKKEPHWLDKWLMLTRKQKGLLSFFFVAVLHAIYSLSPMRR 186
QY 328 SERVFLNMQVQVHANIEVWRIEMVIFSGIMSLGLLSLAVTSIPSVSNALN 387
Db 187 SYRYKLLNWAYQQVQNKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVSDSLT 246
QY 388 WREFSFIQSTGLGVALLISTFHVLIYGWKRAFEFEEYRFTPPNFVLAIVLPSIVIL 444
Db 247 WREFHYIQSKLGIIVSLGLTIIHALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVL 303

RESULT 4
US-09-655-270A-15
Sequence 15, Application US/09655270A
Patent No. 6329151
GENERAL INFORMATION:
APPLICANT: Rouviere, Pierre E.
TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic
FILE REFERENCE: BC1011 US NA
CURRENT APPLICATION NUMBER: US/09/655,270A
CURRENT FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/120,702
PRIOR FILING DATE: 1999-February-19
PRIOR APPLICATION NUMBER: 60/152,542
PRIOR FILING DATE: 1999-September-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Microsoft Office 97
SEQ ID NO 15
LENGTH: 227
TYPE: PRT
ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-655-270A-15

Query Match 6.1%; Score 144.5; DB 4; Length 227;
Best Local Similarity 28.0%; Pred. No. 4e-07;
Matches 60; Conservative 33; Mismatches 88; Indels 33; Gaps 9;
QY 25 IKDARKVTGVIGSGDFAKSLTIRLCYGHVVGISRNPKFASEFFPH-----VVD 75
Db 1 MKSSKIAVVG--GTGPOGKGLAYRFAAGWPVVGISRSARAEAALEVRRRRAGDGA VVS 58
QY 76 VTHHEDALTKTNIIFVAI---HREHYTSLDRLHLLVGKILI-----DVSNMNRIN 123
Db 59 AADNASAAADCPILLVVPYDGHRELVS---ELAPIFAGKLVSCVNPGLGFKSGAYGLD 115
QY 124 QYPESNAEYLASLFPDSLIVKGFNVSA---WALQLGPKDASRQVYICSNNTQARQQVIE 180
Db 116 VEEGSAAEQLRDLVPGATVVAAFHLSAVNLWEHE-GP--LPEDVLVCADDRSKAKDEVAR 172
QY 181 LARQLNFIP-IDGLSLSSAREIENPLRLFTLWR 213
Db 173 LAVAITGRPGIDGALRVAROLEPLTAVLINVR 206

QY 182 ARQLNFIPIDGLSSAREIENPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYAR 241
Db 190 ---FNGTQDELNLSDNNLEELPNDVPHGASGPVILDSRTHSLPSYGLENLKKLRA 246
QY 242 NQOSDFYKPIEIVNKTLP-PIVAITLLSLVY----- 271
Db 247 RSTYNLKKLP-----TLEKLVALMEASLTYPSHCCAFANRRRQISELHPICNKSILRQE 300
QY 272 -----LAGLLAAAYQLYYGTYKRRF----- 305
Db 301 VDMYTQTRGQRSSLAEDNESSYRGFDMTYTEFDYDLCNEVVDVTCSPKPDFAFNPCEDIM 360
QY 306 G-----LLSFFFAMVHA-----YSLCLPMRRSERYLEFLNMAYQQVHANIE 347
Db 361 GYNILRVLIWFISILAITGNIIIVLITTSQYKLTVP-----RFLMCNLAFAD----- 408
QY 348 SWNEEVWRIEMYISFGIMSLGLLSLLAVTSIPSVSNALNWRE-----FSFIQSTL 398
Db 409 -----LCIGIYLLLIASVDIHTKSOYHNYAIDWQTGAGCDAGFTTVFASEL 455
QY 399 G---YVALLISTPH-----VLIYGWKRAFEERYRYFTPTPNFV-LA 435
Db 456 SVYTLTAITLERWHTITHAMQLDCKVQLRHAASVMVMGWIFAFAALPFIIGISSYMKVS 515
QY 436 LVLPISVILDLLQL 449
Db 516 ICLPMDIDSPLSQL 529

RESULT 10
US-09-323-873A-20
; Sequence 20, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Safran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/087,520
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: STEAP-1 PEPTIDE
US-09-323-873A-20

Query Match 4.6%; Score 107; DB 4; Length 34;
Best Local Similarity 58.8%; Pred. No. 0.00019;
Matches 20; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 326 RRSERYFLNMAYQQVHANIEHSWNEEEVWRIEM 359
Db 1 RRSRYKLLNWAYQQQNKEDAWIEHDVWRMEI 34

RESULT 11
US-09-316-083-3
; Sequence 3, Application US/09316083A
; Patent No. 6280942

; GENERAL INFORMATION:
; APPLICANT: The Institute of Physical and Chemical Research
; TITLE OF INVENTION: Endonuclease
; FILE REFERENCE: PH-651
; CURRENT APPLICATION NUMBER: US/09/316,083A
; EARLIER FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: JP98/141861
; EARLIER FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-316-083-3
Query Match 4.3%; Score 101; DB 4; Length 476;
Best Local Similarity 20.7%; Pred. No. 0.047;
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;
QY 62 NPKFASEFFPHVVDVTHHEDALTGTNI-----IFVAIHREHYT-----SLWD 103
Db 254 NPYFVNAFSINI-----KTNLAKEKIFTNIYKLYSDYKINQINNHIPIYNYLK 302
QY 104 LRHLLVGKILIDVSNMNRINQYSPESNAEVLASLF--PDSLIVKGFNVVSAWALQGPKDA 161
Db 303 INNKLPKINIMDIKNYWLAGFTAADGSFLSSMYNPKDTLLFKNM----- 347
QY 162 SRQVYICSNNIQARQQVIELARQLNEFIPIDGLSLSSAREIENPLRLFTLWRGPVVVAIS 221
Db 348 -RPSYVIS-QVETRKELIYLIQE-----SFDL-SISNVKVGNRKLDKFLFTRTTDELMK 400
QY 222 LATFFFLYSFVRDVIHPYARNQOSDFYKIP-----IEIVNKTLPIVAITLLSLVYLAGLLA 277
Db 401 -----FIYF--DKFLPLHDNKQFNKQYIKFRNTFIKSYNNWNRVFGVLVSE--YINNIKI 451
QY 278 AAYQLYYGTYK 288
Db 452 DNYDYYYNKKY 462

RESULT 12
US-08-724-974A-2
; Sequence 2, Application US/08724974A
; Patent No. 5912335
; GENERAL INFORMATION:
; APPLICANT: Derk J. Bergsma, Catherine E. Ellis
; TITLE OF INVENTION: A No. 5912335el G-Protein Coupled Receptor
; TITLE OF INVENTION: HUVCT36
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road, P.O. Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,974A
; FILING DATE: October 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: William T. Han
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50022

QY 122 INQYPESNAEYLSPDLSLIVKGFNVVSAWALQLGPKDASQVYICSNNIQARQQVIEL 181
Db 157 IH-----IVARNFMGLSFESVILWL-----SKNGIEIHNCA----- 189
QY 182 ARQLNFIPIDGLSSAREIENPLRLFTLWRGPPVVAISLATFFFLYSFVRDVIHPYAR 241
Db 190 ---FNGTQDELNLSDNNLEELPNDVFGASGVPILDIS-----RTKVHSLPN 235
QY 242 NQOSDFYKIPFIEIVN--KTLP-----IVAITLLSLVY-----LAGLLAAAYQLY----- 283
Db 236 HGLENKLRARSTYRLKKLPNLDKFTVLMSEASLTYP SHCCAFANLKRQISELHPICNKS 295
QY 284 -----YG-----TKYRRF-----PPWLETWLQC 301
Db 296 ILRQDIDMTQIGDQVRVSLIDDEPSYKGSDDMMYNEFDYDLONEVVDVTCSPKPDAFNPC 355
QY 302 RKQLG-----LLSFFAMVHVA-----YSLCLPMRRRSERYLFLNMAYQQVHA 343
Db 356 EDIMGYNILRVLIWFISILAITGNTTVLVVLTTSQYKLTVP-----RFLMCNLAFAAD--- 407
QY 344 NIENSWNEEEVWRIEMYISFGIMSLGLLSLLAVTSIPSVSNALNWRE-----FSFI 394
Db 408 -----LCIGIYLLIASVDIHTKSOYHNYAIDWQTGAGCDAAGFTTVF 450
QY 395 QSTLG---YVALLISTFH-----VLIYGWKRAFEEEYRYFTPPNF 432
Db 451 ASELSVYTLTAITLERWHTITHAMQLECKVQLRHAASVMVLGWTFAFAAALFPPIFGISSY 510
QY 433 V-LALVLPISIVILDLLQL 449
Db 511 MKVSICLPMDIDSPLSQL 528

RESULT 15

US-08-724-974A-3
; Sequence 3, Application US/08724974A
; Patent No. 5912335
; GENERAL INFORMATION:
; APPLICANT: Derk J. Bergsma, Catherine E. Ellis
; TITLE OF INVENTION: A No. 5912335el G-Protein Coupled Receptor
; TITLE OF INVENTION: HUVCT36
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road, P.O. Box 1539
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,974A
; FILING DATE: October 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: William T. Han
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610 270 5024
; TELEFAX: 610 270 5090
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365
; TYPE: Amino Acid
; TOPOLOGY: Linear

US-08-724-974A-3

Query Match 4.0%; Score 93; DB 2; Length 365;
Best Local Similarity 22.2%; Pred. No. 0.22;
Matches 47; Conservative 37; Mismatches 76; Indels 52; Gaps 9;
QY 255 VNKTLL-PIVAITLLSLVYLAGLLAAAYQIYYGTYRRFPWLETWLCRKQLGLLSFFFA 313
Db 19 IHQTLAPVVVTVL-----VVGFPANCLSLYFG-----YLO--IKARNELGVYLCNLT 64
QY 314 MVHVAYSCLCLPMRRSERYLFLNMAYQQVHANIEANSWNEEEVWRIEMVIFSFGIMSLGLLSL 373
Db 65 VADLFYICSLPFW-----LQYVLQHDNWSHGDLSCQVCVGILLYENIYI-SVGFLCC 114
QY 374 LAVTSIPSVSNALNWREFSFIQSTLGYVA-----LLISTF----- 408
Db 115 ISVDRYLAVAHPPFRFHQFRTLKAARVTVVVIWAKELLTSIYFLMHEEVIEDENQHRVCFE 174
QY 409 HVLIYGWKRAFEEEYRYFTPPNFVLAIVLPS 440
Db 175 HYPIQAWQRAI--NYRFLVGLFPLPICLLLAS 204

Search completed: January 17, 2003, 18:59:36
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 18:58:14 ; Search time 36 Seconds
(without alignments)
2598.483 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGMGSPKSLSETCLPN.....ALVLPISIVILDLLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1316	56.0	488	11	Q99P41	Q99p41 rattus norv
2	1306	55.6	514	11	Q924Z1	Q924z1 mus musculu
3	1269	54.0	488	4	Q9NVB5	Q9nvb5 homo sapien
4	1245.5	53.0	487	4	Q8TF03	Q8tf03 homo sapien
5	1074.5	45.7	456	4	Q8TDP3	Q8tdp3 homo sapien
6	1064.5	45.3	470	11	Q923B6	Q923b6 mus musculu
7	1057.5	45.0	470	11	Q91W31	Q91w31 mus musculu
8	1041.5	44.3	474	11	Q91ZE8	Q91ze8 mus musculu
9	728	31.0	338	6	Q9GL50	Q9gl50 sus scrofa
10	715	30.4	339	11	Q924Z2	Q924z2 mus musculu
11	707	30.1	339	11	Q9CWR7	Q9cwr7 mus musculu
12	704	29.9	339	11	Q924J9	Q924j9 mus musculu
13	628	26.7	283	4	Q9H5R1	Q9h5r1 homo sapien
14	577	24.5	264	4	Q9Y6U5	Q9y6u5 homo sapien
15	541	23.0	283	4	Q8WWB0	Q8wwb0 homo sapien
16	426	18.1	132	11	Q9D5R1	Q9d5r1 mus musculu

17	324	13.8	143	4	Q9H7Y1	Q9h7y1 homo sapien
18	170	7.2	208	16	Q8XQS2	Q8xqs2 ralstonia s
19	160	6.8	198	16	Q930K7	Q930k7 rhizobium m
20	159.5	6.8	211	16	Q8YK44	Q8yk44 anabaena sp
21	156.5	6.7	239	16	Q9RKR6	Q9rkr6 streptomyce
22	151	6.4	200	16	Q98C62	Q98c62 rhizobium l
23	147	6.3	213	17	Q29059	Q29059 archaeoglob
24	144.5	6.1	226	2	Q9AH05	Q9ah05 rhodococcus
25	143.5	6.1	223	17	Q8TYQ9	Q8tyq9 methanopyru
26	142	6.0	212	17	Q29370	Q29370 archaeoglob
27	142	6.0	224	1	O59661	O59661 methanobact
28	137.5	5.8	199	2	Q9F418	Q9f418 mycobacteri
29	136.5	5.8	217	16	Q9JN78	Q9jnn78 streptomyce
30	128.5	5.5	220	2	Q8VPU7	Q8vpu7 nocardioide
31	121	5.1	191	16	Q9CF67	Q9cf67 lactococcus
32	119.5	5.1	216	2	O50637	O50637 leptospira
33	119.5	5.1	254	16	Q8UKE0	Q8uke0 agrobacteri
34	117.5	5.0	429	17	Q970N3	Q970n3 sulfolobus
35	116.5	5.0	242	16	Q9PCP1	Q9pcp1 xylella fas
36	114	4.8	207	16	Q9RIU0	Q9riu0 streptomyce
37	112.5	4.8	207	16	Q9RIT8	Q9rit8 streptomyce
38	112.5	4.8	247	17	Q8TIC0	Q8tic0 methanosarc
39	112	4.8	408	5	Q9V8W8	Q9v8w8 drosophila
40	111.5	4.7	1228	3	Q12754	Q12754 saccharomyc
41	110.5	4.7	222	17	Q9HMC5	Q9hmc5 halobacteri
42	110	4.7	341	17	Q96Y98	Q96y98 sulfolobus
43	109	4.6	219	16	Q9RIU2	Q9riu2 streptomyce
44	109	4.6	320	12	Q9YVM4	Q9yvm4 melanoplus
45	109	4.6	1301	10	O23530	O23530 arabidopsis

ALIGNMENTS

RESULT 1

- Q99P41 ID Q99P41 PRELIMINARY; PRT; 488 AA.
- AC Q99P41;
- DT 01-JUN-2001 (TREMBLrel. 17, Created)
- DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
- DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
- DE Tumor suppressor phyde.
- OS Rattus norvegicus (Rat).
- OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
- OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
- OX NCBI_TaxID=10116;
- RN [1]
- RP SEQUENCE FROM N.A.
- RC STRAIN=DUNNING; TISSUE=PROSTATE CANCER TUMOR;
- RX MEDLINE=20424188; PubMed=10969787;
- RA Steiner M.S., Zhang X., Wang Y., Lu Y.;
- RT "Growth inhibition of prostate cancer by an adenovirus expressing a novel tumor suppressor gene, phyde.";
- RL Cancer Res. 60:4419-4425(2000).
- RN [2]
- RP SEQUENCE FROM N.A.
- RC STRAIN=DUNNING; TISSUE=PROSTATE CANCER TUMOR;
- RA Lu Y., Rinaldy A.R., Steiner M.S.;
- RL Submitted (JAN-2001) to the EMBL/GenBank/DBBJ databases.
- RN [3]
- RP SEQUENCE FROM N.A.
- RA Steiner M.S., Rinaldy A.;
- RT "A novel prostate cancer associated gene.";
- RL Submitted (FEB-2000) to the EMBL/GenBank/DBBJ databases.
- DR EMBL; AF335281; AAK00361.1; -.
- DR EMBL; AF238865; AAL78207.1; -.
- SQ SEQUENCE 488 AA; 54640 MW; 59FF07121919FDCB CRC64;

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 9.9e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGSRNPKF 65

Db	14	LVDSGSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFFVVGSRNPKR	63
QY	66	ASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVLGKILIDVSNMNRIN--	123
Db	64	TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYSSLSCLADQLAGKILVDVSNPTEKERL	123
QY	124	QYPESNAEYLASLFPDSLIVKGFNVVSAWALQLGPKDASROVVICSNNIQARQQVIELAR	183
Db	124	QHRQSNAEYLASLFPACTVVKAFNVISAWALQAGPRDGNRQVLICGDQLEAKHTVSEMAR	183
QY	184	QLNFIPIDLGLSSSAREIENLPLRFTLWRGPVVVAISLATFEFLYSFVRDVIHPYARNQ	243
Db	184	AMGFTPLDMGSLASAREVEAIPRLPWSKVPTLLALGLSTQSYAYNFIRDVLQPYIRKD	243
QY	244	QSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKYYRRFPWPWLETWLOCRK	303
Db	244	ENKEYKMPLSVWNTTIPCVAAYVLLSLVYLPGLVAAALQLRRGTYKQRFPPDWLDHWLQHRK	303
QY	304	QLGLLSFFFAMVHVAYSCLPMRRSERYFLNMAVQQVQHANIENSNWNEEEVWRIEMYISF	363
Db	304	QIGLLSFFFAMLHALYSFCLPLRRSHRYDLVNLAVKQVLANKSRLWVEEVWRMEIYLSL	363
QY	364	GIMSLGLSLLAVTSIPSVSNALNWREFFSIQSTLGYVALLISTFHVLIYGWKRAFEEEY	423
Db	364	GVLAGMLSLLAVTSIPSIANSLNWKEEFSVQSTLGFVALMLSTMHTLTLYGTWTRAFEEH	423
QY	424	YRFYTPPNEVLALVLPISIVIL	444
Db	424	YKFYLPPTFTLTLLPCVIL	444
RESULT 2			
ID	Q92421	PRELIMINARY;	PRT; 514 AA.
AC	Q92421;		
DT	01-DEC-2001 (TremBLrel. 19, Created)		
DT	01-DEC-2001 (TremBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TremBLrel. 21, Last annotation update)		
DE	Dudulin 2.		
GN	1010001D0IRIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Serru V., Lamblin D., Lenoir C., Manivet P., Vaubourdolle M.,		
RA	Kellermann O., Loric S.;		
RT	"Molecular cloning and expression of mouse dudulin 2.";		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY029586; AAK50539.1; -.		
DR	MGD; MGI:1915678; 1010001D0IRIK.		
SQ	SEQUENCE 514 AA; 57268 MW; 339886C288AEC0E2 CRC64;		
Query Match 55.6%; Score 1306; DB 11; Length 514;			
Best Local Similarity 56.2%; Pred. No. 6.6e-96;			
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;			
QY	6	MMGSPKSLSSETCLPNGINGIKDARKVTVGIVSGDFAKSLTIRLIRCGYHVVGSRNPKF	65
Db	14	LVDSGSLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFFVVGSRNPKR	63
QY	66	ASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVLGKILIDVSNMNRIN--	123
Db	64	TAGLFPSLAQVTFQEEAVSSPEVIFVAVFREHYSSLSCLADQLAGKILVDVSNPTEKEHL	123
QY	124	QYPESNAEYLASLFPDSLIVKGFNVVSAWALQLGPKDASROVVICSNNIQARQQVIELAR	183
Db	124	QHRQSNAEYLASLFPACTVVKAFNVISAWALQAVPRDGNRQVLICGNDSKAKQRVMEMAR	183
QY	184	QLNFIPIDLGLSSSAREIENLPLRFTLWRGPVVVAISLATFEFLYSFVRDVIHPYARNQ	243

Db	184	AMGFTPLDMGSLASAREVEAIPRLPWSKVPTLLALGLFVCFTYTNFIRDVLQPYIRKD	243
QY	244	QSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKYYRRFPWPWLETWLOCRK	303
Db	244	ENKEYKMPLSVWNTTIPCVAAYVLLSLVYLPGLVAAALQLRRGTYKQRFPPDWLDHWLQHRK	303
QY	304	QLGLLSFFFAMVHVAYSCLPMRRSERYFLNMAVQQVQHANIENSNWNEEEVWRIEMYISF	363
Db	304	QIGLLSFFFAMLHALYSFCLPLRRSHRYDLVNLAVKQVLANKSRLWVEEVWRMEIYLSL	363
QY	364	GIMSLGLSLLAVTSIPSVSNALNWREFFSIQSTLGYVALLISTFHVLIYGWKRAFEEEY	423
Db	364	GVLAGMLSLLAVTSIPSIANSLNWKEEFSVQSTLGFVALILSTMHTLTLYGTWTRAFEEH	423
QY	424	YRFYTPPNEVLALVLPISIVIL	444
Db	424	YKFYLPPTFTLTLLPCVIL	444
RESULT 3			
ID	Q9NVB5	PRELIMINARY;	PRT; 488 AA.
AC	Q9NVB5;		
DT	01-OCT-2000 (TremBLrel. 15, Created)		
DT	01-OCT-2000 (TremBLrel. 15, Last sequence update)		
DT	01-DEC-2001 (TremBLrel. 19, Last annotation update)		
DE	CDNA FLJ10829 fis, clone NT2RP4001138 (Dudulin 2).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,		
RA	Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,		
RA	Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,		
RA	Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;		
RT	"NEDO human cDNA sequencing project.";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Serru V., Manivet P., Lenoir C., Eschwege P., Lamblin D.,		
RA	Vaubourdolle M., Kellermann O., Loric S.;		
RT	"Dudulin 2, a new tumor antigen expressed in various human tumors.";		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AK001691; BAA91839.1; -.		
DR	EMBL; AY029585; AAK50538.1; -.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.		
SQ	SEQUENCE 488 AA; 54616 MW; BC0BCA483335AAD6 CRC64;		
Query Match 54.0%; Score 1269; DB 4; Length 488;			
Best Local Similarity 54.3%; Pred. No. 5.5e-93;			
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;			
QY	3	SISMMGSPKLSSETCLPNGINGIKDARKVTVGIVSGDFAKSLTIRLIRCGYHVVGSRN	62
Db	11	SLHLVDSDSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGFKVVGSRN	60
QY	63	PKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVLGKILIDVSN--NM	120
Db	61	PKRTARLYPSAAQVTFQEEAVSSPEVIFVAVFREHYSSLSLSDQLAGKILVDVSNPTEQ	120
QY	121	RINQYPESNAEYLASLFPDSLIVKGFNVVSAWALQLGPKDASROVVICSNNIQARQQVIE	180
Db	121	EHLQHRESNAEYLASLFPCTCTVVKAFNVISAWTLQAGPRDGNRQVPICGDQPEAKRAVSE	180
QY	181	LARQLNFIPIDLGLSSSAREIENLPLRFTLWRGPVVVAISLATFEFLYSFVRDVIHPYA	240
Db	181	MALANGFMPVDMGSLASAWEVEAMPLRLLPWKVPTLLALGLFVCFTYAYNEFVRDVLQPYV	240
QY	241	RNQSDFYKIPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYYGTYKYYRRFPWPWLETWLO	300

		Matches	128;	Conservative	51;	Mismatches	63;	Indels	0;	Gaps	0;
QY	209	FTLWRGPVVVAISLATFFFLYSFVRDVIHPYARNQQSDFYKIPIEIVNKTLPPIVAITLLS	268								
Db	68	FPNWRLPVKVAALISLTLFLYTLTREIYPLVTSREQYFYKIPILVINKVLPWVAITLLA	127								
QY	269	LVYLAGLLAAAYQLYGYTKYRRFPPLWLTQCRKQLGLLSFFFFAMVHVAYSCLCLPMRRS	328								
Db	128	LVYLPGEAAVQVQLRNGTKYKKFPPLDRWMLAKQFGLLSFFFAVLHAVYSLSYPMRRS	187								
QY	329	ERYLFNWAYQQVQHANIENSWNEEEVWRIEMYISFGIMSLGLLSLAVTSIPSVSNALNW	388								
Db	188	YRYKLLNWAYKQVQONKEDAWVEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVSDSLTW	247								
QY	389	REFSFIOSTLGYVALLISTFHVLIYGWKRAFEEYRYFTPPNFVIALVLPISIVILDLLQ	448								
Db	248	REFHYIQSKLGIVSLLGTVHALVFAWNKWDVVSQFVWYMPPTFMAVFLPTLVLICKIA	307								
QY	449	LC 450									
Db	308	LC 309									
RESULT 12											
Q924J9											
ID	Q924J9	PRELIMINARY;	PRT;	339	AA.						
AC	Q924J9;										
DT	01-DEC-2001 (TremBLrel. 19, Created)										
DT	01-DEC-2001 (TremBLrel. 19, Last sequence update)										
DT	01-JUN-2002 (TremBLrel. 21, Last annotation update)										
DE	Six-transmembrane epithelial antigen of the prostate.										
GN	STEAP.										
OS	Mus musculus (Mouse).										
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;										
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.										
OX	NCBI_TaxID=10090;										
RN	[1]										
RP	SEQUENCE FROM N.A.										
RC	STRAIN=C57BL/6;										
RX	MEDLINE=21371909; PubMed=11479226;										
RA	Yang D., Holt G.E., Velders M.P., Kwon E.D., Kast W.M.;										
RT	"Murine six-transmembrane epithelial antigen of the prostate, prostate										
RT	stem cell antigen, and prostate-specific membrane antigen: prostate-										
RT	specific cell-surface antigens highly expressed in prostate cancer of										
RT	transgenic adenocarcinoma mouse prostate mice.";										
RL	Cancer Res. 61:5857-5860(2001).										
DR	EMBL; AF297098; AAK83126.1; -.										
DR	MGD; MGI:1917608; Steap.										
KW	Transmembrane.										
SQ	SEQUENCE 339 AA; 39318 MW; 4B26A71FF559E84F CRC64;										
Query Match											
Best Local Similarity 52.9%; Score 704; DB 11; Length 339;											
Matches 128; Conservative 51; Mismatches 63; Indels 0; Gaps 0;											
QY	209	FTLWRGPVVVAISLATFFFLYSFVRDVIHPYARNQQSDFYKIPIEIVNKTLPPIVAITLLS	268								
Db	68	FPNWRLPVKVAALISLTLFLYTLTREIYPLVTSREQYFYKIPILVINKVLPWVAITLLA	127								
QY	269	LVYLAGLLAAAYQLYGYTKYRRFPPLWLTQCRKQLGLLSFFFFAMVHVAYSCLCLPMRRS	328								
Db	128	LVYLPGEAAVQVQLRNGTKYKKFPPLDRWMLARKQFGLLSFFFAVLHAVYSLSYPMRRS	187								
QY	329	ERYLFNWAYQQVQHANIENSWNEEEVWRIEMYISFGIMSLGLLSLAVTSIPSVSNALNW	388								
Db	188	YRYKLLNWAYKQVQONKEDAWVEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVSDSLTW	247								
QY	389	REFSFIOSTLGYVALLISTFHVLIYGWKRAFEEYRYFTPPNFVIALVLPISIVILDLLQ	448								
Db	248	REFHYIQSKLGIVSLLGTVHALVFAWNKWDVVSQFVWYMPPTFMAVFLPTLVLICKIA	307								
QY	449	LC 450									

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 19:00:21 ; Search time 16 Seconds
(without alignments)
564.053 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMMGSPKSLSETCLPN.....ALVLPISIVILDLLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2294	97.6	490	9	US-09-802-520-1 Sequence 1, Appli
2	1269	54.0	488	10	US-09-747-835A-13 Sequence 13, Appl
3	1088	46.3	459	9	US-09-965-529-21 Sequence 21, Appl
4	844	35.9	179	10	US-09-864-761-37654 Sequence 37654, A
5	768	32.7	237	10	US-09-747-835A-15 Sequence 15, Appl
6	736	31.3	141	10	US-09-963-896-1 Sequence 1, Appli
7	717	30.5	267	10	US-09-747-835A-50 Sequence 50, Appl
8	717	30.5	267	10	US-09-747-835A-51 Sequence 51, Appl
9	717	30.5	339	9	US-10-012-896-879 Sequence 879, App
10	717	30.5	339	9	US-09-802-520-11 Sequence 11, Appl
11	717	30.5	339	9	US-09-895-793-879 Sequence 879, App
12	717	30.5	339	9	US-09-895-814-879 Sequence 879, App
13	717	30.5	339	10	US-09-759-143-879 Sequence 879, App
14	717	30.5	339	10	US-09-780-669-879 Sequence 879, App
15	717	30.5	339	10	US-09-822-827-879 Sequence 879, App
16	324	13.8	132	10	US-09-789-561-89 Sequence 89, Appl
17	202	8.6	104	10	US-09-864-761-47234 Sequence 47234, A
18	196	8.3	38	10	US-09-864-761-38670 Sequence 38670, A
19	171.5	7.3	208	9	US-09-738-626-6627 Sequence 6627, Ap

20	107.5	4.6	695	10	US-09-804-626-8	Sequence 8, Appli
21	104	4.4	396	10	US-09-815-242-10164	Sequence 10164, A
22	101	4.3	476	10	US-09-774-414-3	Sequence 3, Appli
23	101	4.3	735	10	US-09-815-242-12089	Sequence 12089, A
24	99	4.2	365	9	US-10-109-533A-2	Sequence 2, Appli
25	97.5	4.1	516	10	US-09-804-626-4	Sequence 4, Appli
26	97	4.1	675	10	US-09-877-804-7	Sequence 7, Appli
27	97	4.1	692	10	US-09-877-804-6	Sequence 6, Appli
28	96	4.1	299	10	US-09-903-456-61	Sequence 61, Appl
29	95	4.0	224	9	US-09-738-626-4894	Sequence 4894, Ap
30	94	4.0	284	10	US-09-903-456-83	Sequence 83, Appl
31	94	4.0	292	10	US-09-903-456-35	Sequence 35, Appl
32	94	4.0	293	10	US-09-903-456-45	Sequence 45, Appl
33	91	3.9	485	9	US-10-055-025-2	Sequence 2, Appli
34	90.5	3.8	465	10	US-09-915-181A-8	Sequence 8, Appli
35	90.5	3.8	1617	9	US-10-090-453A-2	Sequence 2, Appli
36	89.5	3.8	968	10	US-09-753-008-7	Sequence 7, Appli
37	89	3.8	1624	12	US-10-090-454-2	Sequence 2, Appli
38	88.5	3.8	675	10	US-09-733-630-2	Sequence 2, Appli
39	88	3.7	387	8	US-08-945-749-8	Sequence 8, Appli
40	87.5	3.7	299	10	US-09-903-456-66	Sequence 66, Appl
41	87.5	3.7	675	9	US-09-928-530-2	Sequence 2, Appli
42	87	3.7	348	10	US-09-877-804-8	Sequence 8, Appli
43	87	3.7	467	10	US-09-991-212A-3	Sequence 3, Appli
44	87	3.7	801	10	US-09-815-242-5518	Sequence 5518, Ap
45	87	3.7	801	10	US-09-815-242-12519	Sequence 12519, A

ALIGNMENTS

RESULT 1
US-09-802-520-1
; Sequence 1, Application US/09802520
; Publication No. US20020187472A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Faris, Mary
; APPLICANT: Chen, Hwei-Mei
; APPLICANT: Ison, Craig H.
; TITLE OF INVENTION: STEAP-RELATED PROTEIN
; FILE REFERENCE: PC-0037 US
; CURRENT APPLICATION NUMBER: US/09/802,520
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020187472A1 7492448CD1
US-09-802-520-1

Query Match	97.6%	Score 2294;	DB 9;	Length 490;
Best Local Similarity	100.0%;	Pred. No. 1.1e-205;		
Matches 444;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS	60	
Db	1	MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVVIGS	60	
QY	61	RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNM	120	
Db	61	RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNM	120	
QY	121	RINQYPESNAEYLASLFPDSLIVKGFNVVSAAWLQPKDASRQVYICSNNIARQQVIE	180	
Db	121	RINQYPESNAEYLASLFPDSLIVKGFNVVSAAWLQPKDASRQVYICSNNIARQQVIE	180	
QY	181	LARQLNFIPIDLGLSSAREIENPLRLFTLWRGPVVVAISLATFFFLYSFVRDVIHPYA	240	

Db 181 LARQLNFIPIDGLSSLSSAREIENPLRLFTLWGRPVVVAISLATEFFFLYSEVRDVIHPYA 240
QY 241 RNQOSDFYKPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWPWLETWLQ 300
Db 241 RNQOSDFYKPIEIVNKTLPPIVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWPWLETWLQ 300
QY 301 CRKQLGLLSFFFAMVHVAYSCLCPMRRSERYLFNLMAYQQVHANIEWSNNEEEVWRIEMY 360
Db 301 CRKQLGLLSFFFAMVHVAYSCLCPMRRSERYLFNLMAYQQVHANIEWSNNEEEVWRIEMY 360
QY 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWREFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
Db 361 ISFGIMSLGLLSLLAVTSIPSVSNALNWREFSFIQSTLGYVALLISTFHVLIYGWKRAFE 420
QY 421 EYRYRFTPPNFVLALVLPISIVIL 444
Db 421 EYRYRFTPPNFVLALVLPISIVIL 444

RESULT 2

US-09-747-835A-13
; Sequence 13, Application US/09747835A
; Patent No. US20020146692A1

GENERAL INFORMATION:

; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES

; FILE REFERENCE: HYS-37CIP

; CURRENT APPLICATION NUMBER: US/09/747,835A

; PRIOR APPLICATION NUMBER: 2002-03-08

; PRIOR FILING DATE: 2000-12-04

; PRIOR APPLICATION NUMBER: US 09/729,739

; PRIOR FILING DATE: 2000-08-31

; PRIOR APPLICATION NUMBER: US 09/620,312

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: US 09/598,042

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: US 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: US 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13

; LENGTH: 488

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-747-835A-13

Query Match 54.0%; Score 1269; DB 10; Length 488;
Best Local Similarity 54.3%; Pred. No. 2.8e-110;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;

QY 3 SISMMGSPKLSLSECTLPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVHVGSRN 62

Db 11 SLHLVDSDSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGSGFKVVGSRN 60

QY 63 PKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSN--NM 120

Db 61 PKRTARLYPSAAQVTFQEEAVSSPEVIFVAFVREHYSSLCSLSDQLAGKILVDVSNPTEQ 120

QY 121 RINQYPESNAEYLASLFPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQQVIE 180

Db 121 EHLQHRRESNAEYLASLFPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQQVIE 180

RESULT 3

US-09-965-529-21

; Sequence 21, Application US/09965529

; Publication No. US20020182671A1

GENERAL INFORMATION:

; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra

; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS

; FILE REFERENCE: PF-0731 USA

; CURRENT APPLICATION NUMBER: US/09/965,529

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315

; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14

; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: PERL Program

; SEQ ID NO 21

; LENGTH: 459

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20020182671A1 1859305CD1

US-09-965-529-21

Query Match 46.3%; Score 1088; DB 9; Length 459;
Best Local Similarity 46.9%; Pred. No. 1.8e-93;
Matches 204; Conservative 89; Mismatches 136; Indels 6; Gaps 3;

QY 13 LSETC--LPNGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHVHVGSRNPKFASEFF 69

Db 1 MEKTCIDALPLTMN--SSEKQETVCIFGTGDFGRSLGLKMLQCGYSVVFGSRNPQ-KTTL 57

QY 70 FPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDVSNMNRINQYPESN 129

Db 58 LPSGAEVLSYSEAAKKSDDIIIAIHREHYDFTELTEVLNGKILVDISNNLKINQYPESN 117

QY 130 AEYLASLFPDLSLVKGFNVSAWALQGPKDASRQVYICSNNIQARQQVIELARQNLNFP 189

Db 118 AEYLAHLVPGAHVVKAFTISAWALQSGALDASRQVFCGNDKAKQKRVMDIVRNGLGTP 177

QY 190 IDGLSSLSSAREIENPLRLFTLWGRPVVVAISLATEFFFLYSEVRDVIHPYARNQOSDFYK 249

Db 178 MDQGSMAAKEIEKYPLQLFPMWRFPFYLAVLCVFLFFCYCVRDVIYPVYVEKKDNTFR 237

Qy	250	IPIEIVNKTLP	IVAITLLSLVYL	AGLLAAAYQLY	GTYKVRFP	PWLETWLQCR	KQLGLLS	309
		:	:	:	:	:	:	
Dd	238	MAISIPNRFP	TALTLLALVY	LPGVIAAILQ	LYRGTYKVRFP	DWLDDHWMCL	RKQLGLVA	297
		:	:	:	:	:	:	
Qy	310	FFFAMVHVAYS	LCPLMRRSER	VYLEFLNMAY	QQVQHANIENS	WNNEEVWRIEM	YSISFGIMSLG	369
		:	:	:	:	:	:	
Dd	298	LGFAFLHVL	YTLVIPIRYY	VWRRLGNLT	VTQAILKKEN	FSTSSAWLSDS	YVALGILGFF	357
		:	:	:	:	:	:	
Qy	370	LLSLLAVTSP	SVSNALNWRE	SFIQSTGLGY	VALLISTPHVLI	YGWKRAFE	EYYRYFTP	429
		:	:	:	:	:	:	
Dd	358	LFVLLGTSLP	SVSNVAVNW	REFRVQSKGL	YLTILCTAHTL	VYGGKRFL	SPSNLRWYP	417
		:	:	:	:	:	:	
Qy	430	PNFVLALVP	SIVIL	444				
		:	:	:	:	:	:	
Dd	418	AAYVLGLI	PCTVLV	432				
		:	:	:	:	:	:	

RESULT 4
US-09-864-761-37654
; Sequence 37654, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```
; FEATURE:
; OTHER INFORMATION: MAP TO AC002064.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P23523, EVALUAE 8.00e-01
; OTHER INFORMATION: EST_HUMAN HIT: BE783685.1, EVALUAE 9.00e-37
DU-09-864-761-37654
```

	Query Match	35.9%;	Score 844;	DB 10;	Length 179;
	Best Local Similarity	98.8%;	Pred. No. 2.5e-71;		
	Matches 165;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps	0;
QY	1	MESISMMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRCGYHVA	IGS 60		
Db	12	MESISMMGSPKSLSETFLPNGINGIKDARKVTGVIGSGDFAKSLTIRLRCGYHVA	IGS 71		
QY	61	RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDV	SNNM 120		
Db	72	RNPKFASEFFPHVVDVTHHEDALTKTNIIFVAIHREHYTSLWDLRHLVKGKILIDV	SNNM 131		
QY	121	RINQYPESNAEYLASLFPDLSLVKGFNVVSAWALQLGPKDASRQVYI	167		
Db	132	RINQYPESNAEYLASLFPDLSLVKGFNVVSAWALQLGPKDASRQVCI	178		

```

RESULT 5
US-09-747-835A-15
; Sequence 15, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN
; TITLE OF INVENTION: LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-747-835A-15

```

Query Match 32.7%; Score 768; DB 10; Length 237;


```
Query Match      30.5%; Score 717; DB 9; Length 339;
Best Local Similarity 54.9%; Pred. No. 4.1e-59;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLWRGPVVVAISLATEFFFLYSFVRDVIHPYARNQQSDFYKIPIEIVNKTLPVIAITLL 267
Db 67 LFPQWHLPIKIAAIIASLTFLYTLRLREVHPLATSHQQYFYKIPILVINKVLPWVSITLL 126

QY 268 SLVYLAGLLAAAYQLYGYTKYRRFPPLWLTWLCRQKQLGLLSFFFAMVHVAYSCLCPMRR 327
Db 127 ALVYLPGVIAAIVQLHNGTKYKKFPHWLDKWMLTRKQFGLLSFFFAVLHAIYSLSPMRR 186

QY 328 SERYFLNMAYQQVHANIEWSNEEEVWRIEMYSISFGIMSLGLLSLLAVTSIPSVSNALN 387
Db 187 SYRYKLLNWAYQQVQONKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVSDSLT 246

QY 388 WREFSFIQSTLGYVALLISTFHVLIYGWKRAFEEEYRYFTPPNFVLAIVLPSIVIL 444
Db 247 WREFHYIQSKLGIVSLLGTIHALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVLI 303

RESULT 11
US-09-895-793-879
; Sequence 879, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-793-879

Query Match      30.5%; Score 717; DB 9; Length 339;
Best Local Similarity 54.9%; Pred. No. 4.1e-59;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLWRGPVVVAISLATEFFFLYSFVRDVIHPYARNQQSDFYKIPIEIVNKTLPVIAITLL 267
Db 67 LFPQWHLPIKIAAIIASLTFLYTLRLREVHPLATSHQQYFYKIPILVINKVLPWVSITLL 126

QY 268 SLVYLAGLLAAAYQLYGYTKYRRFPPLWLTWLCRQKQLGLLSFFFAMVHVAYSCLCPMRR 327
Db 127 ALVYLPGVIAAIVQLHNGTKYKKFPHWLDKWMLTRKQFGLLSFFFAVLHAIYSLSPMRR 186

RESULT 13
US-09-759-143-879
; Sequence 879, Application US/09759143
```

```
QY 328 SERYFLNMAYQQVHANIEWSNEEEVWRIEMYSISFGIMSLGLLSLLAVTSIPSVSNALN 387
Db 187 SYRYKLLNWAYQQVQONKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVSDSLT 246

QY 388 WREFSFIQSTLGYVALLISTFHVLIYGWKRAFEEEYRYFTPPNFVLAIVLPSIVIL 444
Db 247 WREFHYIQSKLGIVSLLGTIHALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVLI 303

RESULT 12
US-09-895-814-879
; Sequence 879, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-895-814-879

Query Match      30.5%; Score 717; DB 9; Length 339;
Best Local Similarity 54.9%; Pred. No. 4.1e-59;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLWRGPVVVAISLATEFFFLYSFVRDVIHPYARNQQSDFYKIPIEIVNKTLPVIAITLL 267
Db 67 LFPQWHLPIKIAAIIASLTFLYTLRLREVHPLATSHQQYFYKIPILVINKVLPWVSITLL 126

QY 268 SLVYLAGLLAAAYQLYGYTKYRRFPPLWLTWLCRQKQLGLLSFFFAMVHVAYSCLCPMRR 327
Db 127 ALVYLPGVIAAIVQLHNGTKYKKFPHWLDKWMLTRKQFGLLSFFFAVLHAIYSLSPMRR 186

QY 328 SERYFLNMAYQQVHANIEWSNEEEVWRIEMYSISFGIMSLGLLSLLAVTSIPSVSNALN 387
Db 187 SYRYKLLNWAYQQVQONKEDAWIEHDVWRMEIYVSLGIVGLAILALLAVTSIPSVSDSLT 246

QY 388 WREFSFIQSTLGYVALLISTFHVLIYGWKRAFEEEYRYFTPPNFVLAIVLPSIVIL 444
Db 247 WREFHYIQSKLGIVSLLGTIHALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVLI 303

RESULT 13
US-09-759-143-879
; Sequence 879, Application US/09759143
```


Qy	388	WREFSIOSTLGYVALLISTFHVLIYGWKRAFEEYYRFTPTPNFVLAFLVLPISVIL	444
Db	247	WREFHYIQSKLGIVSLLGTIHALIFAWKNWIDIKQFVWYTPPTFEMIAVFLPIVVLI	303

Search completed: January 17, 2003, 19:04:43
Job time : 18 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 18:58:14 ; Search time 14 Seconds
(without alignments)
1345.018 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMMGSPKSLSETCLPN.....ALVLPISIVILDLLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			DB ID	Description
	Score	Match	Length		
1	717	30.5	339	1 STEA_HUMAN	Q9uhe8 homo sapien
2	133	5.7	223	1 F4RE_METJA	Q58896 methanococc
3	120.5	5.1	232	1 F4RE_METTH	O26350 methanobact
4	111.5	4.7	695	1 FSHR_MACFA	P32212 macaca fasc
5	109.5	4.7	206	1 Y538_PASMU	Q9cn97 pasteurella
6	109	4.6	693	1 FSHR_CHICK	P79763 gallus gall
7	109	4.6	712	1 FRE6_YEAST	Q12473 saccharomyc
8	107.5	4.6	629	1 FRE7_YEAST	Q12333 saccharomyc
9	107.5	4.6	695	1 FSHR_HUMAN	P23945 homo sapien
10	104.5	4.4	695	1 FSHR_PIG	P49059 sus scrofa
11	104	4.4	396	1 SOTB_ECO57	P58529 escherichia
12	104	4.4	396	1 SOTB_ECOLI	P31122 escherichia
13	103	4.4	465	1 NPT1_MOUSE	Q61983 mus muscula
14	102	4.3	220	1 Y304_BRUME	Q8yd73 brucella me
15	100	4.3	476	1 RF3_SACBA	P05512 saccharomyc
16	99.5	4.2	452	1 NU4M_BRALA	O79421 branchiosto
17	99	4.2	365	1 GP68_HUMAN	Q15743 homo sapien
18	98	4.2	3411	1 POLG_YEFV1	P03314 y genome po
19	98	4.2	3411	1 POLG_YEFV2	P19901 y genome po
20	97.5	4.1	268	1 YD49_AQUAE	O67364 aquifex aeo
21	97.5	4.1	695	1 FSHR_BOVIN	P35376 bos taurus
22	97	4.1	692	1 FSHR_RAT	P20395 rattus norv
23	96.5	4.1	686	1 FRE1_YEAST	P32791 saccharomyc
24	96	4.1	261	1 PHSC_ECOLI	P77409 escherichia
25	95	4.0	202	1 YAJ1_PSEAE	Q9hva5 pseudomonas
26	95	4.0	393	1 NUCC_SPIOL	P56753 arabidopsis
27	95	4.0	393	1 FSHR_MOUSE	Q9m3i5 spinacia ol
28	95	4.0	692	1 SECY_CYAPA	P25014 cyanophora
29	94.5	4.0	492	1 FSHR_SHEEP	P35379 ovis aries
30	94.5	4.0	695	1 FSHR_BRAFL	O47423 branchiosto
31	93.5	4.0	452	1 GSHR_PLAF7	O15770 plasmodium
32	93.5	4.0	499	1 VLI_REOVD	P15024 reovirus (t
33	93.5	4.0	1233	1	

34	93	4.0	388	1 HMC3_DESVH	P33390 desulfovibr
35	93	4.0	459	1 NU4M_BOVIN	P03910 bos taurus
36	93	4.0	521	1 YT25_CAEEL	Q10934 caenorhabdi
37	92.5	3.9	490	1 GTR1_CHICK	P46896 gallus gall
38	92.5	3.9	500	1 NU4C_OENHO	P58419 oenothera h
39	92.5	3.9	883	1 YHL6_YEAST	P38781 saccharomyc
40	92.5	3.9	1769	1 YJK9_YEAST	P42945 saccharomyc
41	92.5	3.9	2037	1 FAS1_CANAL	P34731 c fatty aci
42	91.5	3.9	495	1 TPS1_YEAST	Q00764 s alpha,alp
43	91.5	3.9	1694	1 CLH_DICDI	P25870 dictyosteli
44	91	3.9	241	1 Y705_CAMJE	Q9ppk2 campylobact
45	91	3.9	409	1 HEM1_THEAC	Q9hkr0 thermoplasm

ALIGNMENTS

RESULT 1

STEAL_HUMAN					
ID	STEAL_HUMAN	STANDARD;	PRT;	339	AA.
AC	Q9UHE8; O95034;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Six transmembrane epithelial antigen of prostate.				
GN	STEAP OR STEAP1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20056277; PubMed=10588738;				
RA	Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,				
RA	Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,				
RA	Jakobovits A., Saffran D.C., Afar D.E.H.;				
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in				
RT	human prostate tumors.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;				
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBDJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Skin;				
RA	Strausberg R.;				
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBDJ databases.				
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PROSTATE TUMORS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF186249; AAF17479.1; -.				
DR	EMBL; AC005053; AAC79150.1; ALT_INIT.				
DR	EMBL; AC004969; AAD15620.1; ALT_INIT.				
DR	EMBL; BC011802; AAH11802.1; -.				
DR	Genew; HGNC:11378; STEAP.				
DR	MIM; 604415; -.				
KW	Transmembrane; Antigen.				
FT	TRANSMEM 71 91				POTENTIAL.
FT	TRANSMEM 119 139				POTENTIAL.
FT	TRANSMEM 164 184				POTENTIAL.
FT	TRANSMEM 218 238				POTENTIAL.
FT	TRANSMEM 258 278				POTENTIAL.
FT	TRANSMEM 291 311				POTENTIAL.
SQ	SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;				


```
RESULT 5
Y538_PASMU
ID Y538_PASMU STANDARD; PRT; 206 AA.
AC Q9CN97;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0191 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006089; AAK02622.1; -
DR Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 7 26 POTENTIAL.
FT TRANSMEM 41 63 POTENTIAL.
FT TRANSMEM 76 98 POTENTIAL.
FT TRANSMEM 113 135 POTENTIAL.
FT TRANSMEM 142 164 POTENTIAL.
FT TRANSMEM 169 191 POTENTIAL.
SQ SEQUENCE 206 AA; 23653 MW; 2F277729D7F11194 CRC64;

Query Match 4.7%; Score 109.5; DB 1; Length 206;
Best Local Similarity 23.8%; Pred. No. 0.44;
Matches 56; Conservative 44; Mismatches 72; Indels 63; Gaps 13;

QY 228 LYSEVRDVIH-----PYA-----RNQSDFYKIPIEIVNKTLPIVAITLLSLVLAGL 275
Db : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
1 MLSLFRITIIHVCCGLGPVAVLAWVLLSGDESQLGADPIKETQHFGLGFSALTILLIMFILG- 59

QY 276 LAAAYQLYGTGYRRFPWPWLETWLQCRKQLGLLSFFFAVHVHVAYSCLPLMRRSERVFLN 335
Db : : : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
60 -----KVFYLLK----QPQLQV---LRRALGLWAFVYVLHV-YA-----YLALE 96

QY 336 MAYQQVHANISNWEVEEVRNEMYSIFGIMSLGLLSLLAVTSIPSVSNALNWRFSFIQ 395
Db : | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
97 LGYD-----FSLFVQELVNR--GYLIIGAIAFLILTLMALSSWSYLK--LKMKGWVFWYL 146

QY 396 STLGYVALLISTFHVLIYGVKRAFEVEEYRFPNPFVLAALVLPISIVILDLLQLC 450
Db ||| ||| : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
147 HQLGGYALLLGAIH---YVW-----SVKNVTFSSML--YLILSIMILC 184

RESULT 6
FSHR_CHICK
ID FSHR_CHICK STANDARD; PRT; 693 AA.
AC P79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin
DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=93323357;
RA Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;
RT "The cDNA cloning and transient expression of a chicken gene encoding
RT a follicle-stimulating hormone receptor.";
RL Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgham J.T., Foster D.N., Johnson A.L.;
RT "Characterization of the chicken follicle-stimulating hormone
RT receptor (cFSH-R) complementary deoxyribonucleic acid, and expression
RT of cFSH-R messenger ribonucleic acid in the ovary.";
RL Biol. Reprod. 55:1055-1062(1996).
CC -!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC FSH/LSH/TSH SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D87871; BAA13487.1; -
DR EMBL; U51097; AAC60030.1; -
DR HSSP; P23945; 1XUN.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 693 FOLLICLE STIMULATING HORMONE RECEPTOR.
FT DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 1 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 421 2 (POTENTIAL).
FT DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 444 465 3 (POTENTIAL).
FT DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 486 508 4 (POTENTIAL).
FT DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 529 550 5 (POTENTIAL).
FT DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 574 597 6 (POTENTIAL).
FT DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 609 630 7 (POTENTIAL).
FT DOMAIN 631 693 EXTRACELLULAR (POTENTIAL).
FT REPEAT 45 68 CYTOPLASMIC (POTENTIAL).
FT REPEAT 69 93 LRR 1.
FT REPEAT 95 118 LRR 2.
FT REPEAT 119 143 LRR 3.
FT REPEAT 168 192 LRR 4.
FT REPEAT 192 192 LRR 5.
```

FT	REPEAT	193	216	LRR 6.
FT	REPEAT	218	240	LRR 7.
FT	DISULFID	442	517	BY SIMILARITY.
FT	CARBOHYD	47	47	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	268	268	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	4	4	G -> D (IN REF. 2).
FT	CONFLICT	88	88	G -> A (IN REF. 2).
FT	CONFLICT	140	140	K -> R (IN REF. 2).
FT	CONFLICT	174	174	I -> T (IN REF. 2).
FT	CONFLICT	191	191	N -> S (IN REF. 2).
FT	CONFLICT	329	329	V -> L (IN REF. 2).
SQ	SEQUENCE	693 AA;	78697 MW;	46F98699635A1BEC CRC64;
Query Match 4.6%; Score 109; DB 1; Length 693;				
Best Local Similarity 17.2%; Pred. No. 2;				
Matches 100; Conservative 72; Mismatches 173; Indels 238; Gaps 21;				
QY	18	LPNG-INGIKDARKVTVG-----VIGSGDFAKSLTIRLCRGYHVVGISRNPKFASEFFP	71	
Db	61	IPKGAFTHLHLEKIEISQNDALIEIEGNVFSSL-----PKL-----	97	
QY	72	HVVDVTHHEDALTKNIIIFVAIHREHYTSLWDLRHLVGLV-----KILI	114	
Db	98	-----HEIRIEKANKL-MKIDQDAFOHPLSLRYLLISNTGLSFLPVVHKVHSFQKVLL	149	
QY	115	DVSNMNRINQYPESNAEYLASLPDPSLIVKGFNVVSAWALQGLPKDASRQVYICSNNIQA	174	
Db	150	DVQDNHIRTIERNTFMGLSS---ESVILR-----LNKNGIQE	184	
QY	175	RQOVIELARQLNFIPIDLGLSSAREIENPLRLFTLWRGPVWVAISLATEFFFLYSFVRD	234	
Db	185	IKD-----HAFNGTCLDELNSDYNLEKLPKVFQGAIGPVVLDISRTISFLPSHGLE	239	
QY	235	VIHPYARNQSDFYKIPIEIVNKTLPPIVAITLLSLVLAGL-----	275	
Db	240	FIKKLRARSTYKLLKLP--DVNKFRLIEANFTYPSHCCAFNRKNTQNTTEFYPICSMSPA	297	
QY	276	-----LAAAYQLYGYTKY---RRF-----PPWLETWLQCR	302	
Db	298	KQDLGEQTGKRKHRRSAAEDYISHYGTFRGFPVENEFDYGLCNEVDVFCSPKPDAPNPCE	357	
QY	303	KQLG-----LLSFFFMVHVA-----YSLCLPMRRSERYLFLNMAVQOVHAN	344	
Db	358	DIMGYNVLRVLWIFINILATGNTTVLIILISSQYKLTVP-----RFLMCNLAFA	408	
QY	345	IENSWNEEVWRIEMYISFGIMSLGLLSLLAVTSIPSVSNALNWRE-----PSFIQ	395	
Db	409	-----LCIGIYLLFIASVDIQTKSRYYNVAIDWQTGAGCNAAGFTVFA	452	
QY	396	STLGYVALLISTFH-----VLIYGWKRAF-----	419	
Db	453	SELSVYTLTVITLERWHTITTYAMQLNRKVRRLRAVIIMVFGWMFAFTVALLPIFGISSYM	512	
QY	420	-----EEYYRFFYTPPNFVLALVLPISIVILDLLQLCRY	452	
Db	513	KVSICLPMHIETPFSQAYV--IFLLVNLVLAFAVIICICICIY	553	
RESULT 7				
FRE6_YEAST				
ID	FRE6_YEAST	STANDARD;	PRT;	712 AA.
AC	Q12473;			
DT	15-DEC-1998	(Rel. 37, Created)		
DT	15-DEC-1998	(Rel. 37, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Ferric reductase	transmembrane component 6 precursor	(EC 1.6.99.13)	
DE	(Ferric-chelate	reductase 6).		
GN	FRE6	OR YLL051C	OR L0593.	
OS	Saccharomyces cerevisiae	(Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota;	Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales;	Saccharomycetaceae; Saccharomyces.		

OX	NCBI_TaxID=4932;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=S288c;
RA	Wedler H., Wambutt R.;
RL	Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC	-!- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC	-!- COFACTOR: FAD (PROBABLE).
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC	-!- SIMILARITY: BELONGS TO THE FRE / CYBB FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; Z47973; CAA88006.1; -.
DR	EMBL; Z73156; CAA97503.1; -.
DR	SGD; S0003974; FRE6.
DR	InterPro; IPR002916; Ferric_reduct.
DR	Pfam; PF01794; Ferric_reduct; 1.
KW	Oxidoreductase; Electron transport; Transmembrane; Iron transport;
KW	FAD; NAD; Glycoprotein; Signal; Multigene family.
FT	SIGNAL 1 17 POTENTIAL.
FT	CHAIN 18 712 FERRIC REDUCTASE TRANSMEMBRANE COMPONENT
FT	6.
FT	NP_BIND 493 499 FAD (POTENTIAL).
FT	TRANSMEM 168 188 POTENTIAL.
FT	TRANSMEM 245 265 POTENTIAL.
FT	TRANSMEM 288 308 POTENTIAL.
FT	TRANSMEM 329 349 POTENTIAL.
FT	TRANSMEM 361 381 POTENTIAL.
FT	TRANSMEM 388 408 POTENTIAL.
FT	TRANSMEM 417 437 POTENTIAL.
FT	TRANSMEM 478 498 POTENTIAL.
FT	TRANSMEM 550 570 POTENTIAL.
FT	DOMAIN 553 556 POLY-LEU.
FT	CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 712 AA; 81989 MW; 5224F12B51544BAA CRC64;
Query Match 4.6%; Score 109; DB 1; Length 712;	
Best Local Similarity 18.8%; Pred. No. 2;	
Matches 72; Conservative 51; Mismatches 119; Indels 142; Gaps 17;	
QY	88 IIFVAIHREHYTSL-----WDLRHLVLGKILIDVSNMNRINQYPESNAEYLASL 136
Db	183 ITAVFFHMSHYNGLNRLAFASRFVNYIRGHFVLPTFLVD---KHANHFKFLNVEVFTGL 238
QY	137 FPDSLIVKGFNVVSAWALQGLPKDASRQVYICSNNIARQOQVIELARQLNFIPIDLGSL 196
Db	239 MPNSL-----EAW----- 246
QY	197 SAREIENLPLRLFTLWRGPVWVAISLATFFFL-YSFVRDVIHPYARNQQSDFYKIPIEV 255
Db	247 -----IIFGYTLANIIFLSISYIIDPYNLIFNSHLSQFTRL---LA 284
QY	256 NKTLPIVAITLLSLVLAGLLAAAYQLYGYTKYRRFPWPWLETWLQCRKQLGLSFFPAMV 315
Db	285 DRS-GILAFQTQPLIIIFTARNSFLEFLTGKFNFSF-----ISFHKWIGRIMVLNATI 336
QY	316 H-VAYSICLPMRRSERYLFLNMAVQOVHANIENSWNEEVWRIEMYISFGIMSLGLSL 374
Db	337 HSLSYSL-----FAIINHAFFK-----ISNK-----QLYWKFGIASITVLCVL 373
QY	375 AVTSIPSVSNALNWREFSFIQSTLGYVALLI---STFHVLIY-GWKRAFE-----EE 422

Db 374 LVLSLGIVRK---RHYEFFLYTHIILALLFFYCCWQHVKIFNGWKWIVVSLLIWGLEK 429

QY 423 YRYFYTPNFFVLALVLPISVILD 446

Db 430 LFRWN----ILQFRFPKATLINL 449

RESULT 8

FRE7_YEAST

ID FRE7_YEAST STANDARD; PRT; 629 AA.

AC Q12333;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ferric reductase transmembrane component 7 (EC 1.6.99.13) (Ferric-chelate reductase 7).

GN FRE7 OR YOL152W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / FY1679;

RX MEDLINE=96132030; PubMed=8553699;

RA Casamayor A., Aldea M., Casas C., Herrero E., Gamo F.J., Lafuente M.J., Gancedo C., Arino J.;

RA "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast chromosome XV containing seven new open reading frames.";

RT yeast 11-1281-1288(1995).

RL CC -|- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).

CC -|- COFACTOR: FAD (PROBABLE).

CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -|- SIMILARITY: BELONGS TO THE FRE / CYBB FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensed@isb-sib.ch).

CC -----

DR EMBL; 248239; CAA88276.1; -.

DR EMBL; 274894; CAA99174.1; -.

DR SGD; S0005512; FRE7.

DR InterPro; IPR002916; Ferric_reduct.

DR Pfam; PF01794; Ferric_reduct; 1.

KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;

KW FAD; NAD; Glycoprotein; Multigene family.

FT NP_BIND 369 375 FAD (POTENTIAL).

FT TRANSMEM 46 66 POTENTIAL.

FT TRANSMEM 108 128 POTENTIAL.

FT TRANSMEM 168 188 POTENTIAL.

FT TRANSMEM 195 215 POTENTIAL.

FT TRANSMEM 238 258 POTENTIAL.

FT TRANSMEM 266 286 POTENTIAL.

FT TRANSMEM 293 313 POTENTIAL.

FT TRANSMEM 422 442 POTENTIAL.

FT CARBOHYD 330 330 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 541 541 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 629 AA; 71996 MW; 2384480E9289C16F CRC64;

Query Match 4.6%; Score 107.5; DB 1; Length 629;

Best Local Similarity 21.7%; Pred. No. 2.2;

Matches 57; Conservative 46; Mismatches 77; Indels 83; Gaps 17;

QY 214 GPVVVAISLATFFFLYSFVRDVHPYARNQQSDFYKPIEIVNKTLPDIVAITLLSLVY-L 272

Db 123 GTFLVVMATTLYTLLYCFVP---HPFYR-PCAGFGSPPLSV---RAGIMASLVPFVFSL 175

QY 273 AG-----LLAAAYQLYGTGYKRYRFPWPWLETWLQCRKQLGLLSFFFAFMVHVAYSCLLPM 325

Db 176 SGKINVGWLVGLSYE----KINIYHQW-----ASILCFFSWVHV-----IPF 215

QY 326 RRSERYLFLNMAYQQVQHANIENSNNEEEVWRIEMYISFGIMSLGLLSLLAVTSIP----- 380

Db 216 LRQARH---EGGYERMH---QRWKASDMWR-----SGVPPILFLNLLWLSSLPIARRH 262

QY 381 --SVSNALNWREFSFIQSTLGYVALLISTFHV-----LIYGWKRAFEEEYYR- 425

Db 263 FYEIFLQLHW-----ILAVGFYISLF---YHVPPELNHMYLVATIVVW---FAQLFYRL 311

QY 426 ----FYTPPNFVLALVLPISVIL 444

Db 312 AVKGYLRPGRSFMASTIANVSIV 334

RESULT 9

FSHR_HUMAN

ID FSHR_HUMAN STANDARD; PRT; 695 AA.

AC P23945;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Follicle stimulating hormone receptor precursor (FSH-R) (Follitropin receptor).

DE FSHR.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Ovary;

RX MEDLINE=91222171; PubMed=1709010;

RA Minegishi T., Nakamura K., Takakura Y., Ibuki Y., Igarashi M.;

RT "Cloning and sequencing of human FSH receptor cDNA.";

RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=93246012; PubMed=1301382;

RA Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L., Rosenthal J.L., Overton S.A., Wands G.D., Kuzeja J.B., Luchette C.A., Chappel S.C.;

RT "The cloning of the human follicle stimulating hormone receptor and its expression in COS-7, CHO, and Y-1 cells.";

RL Mol. Cell. Endocrinol. 89:141-151(1992).

RN [3]

RP SEQUENCE FROM N.A.

RA Tilly L.T., Aihara T., Nishimori K., Jai X.-C., Billig H., Kowalski K.I., Perlas E.A., Hsueh A.J.;

RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 1-342 FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=93075197; PubMed=1359889;

RA Gromoll J., Gudermann T., Nieschlag E.;

RT "Molecular cloning of a truncated isoform of the human follicle stimulating hormone receptor.";

RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).

RN [5]

RP SEQUENCE OF 1-51 FROM N.A.

RX MEDLINE=95011044; PubMed=7926278;

RA Gromoll J., Dankbar B., Gudermann T.;

RT "Characterization of the 5' flanking region of the human follicle stimulating hormone receptor gene.";

RL Mol. Cell. Endocrinol. 102:93-102(1994).

RN [6]

RP 3D-STRUCTURE MODELING OF 49-228.

RX MEDLINE=96363672; PubMed=8747461;

RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H., Hendrickson W.A., el Tayar N.;

RT "Structural predictions for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions.";

la Barbera A.R.;
"Porcine follicle-stimulating hormone receptor."
Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: RECEPTOR FOR FOLLICLE STIMULATING HORMONE. THE ACTIVITY
OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
ADENYLATE CYCLASE.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
FSH/LSH/TSH SUBFAMILY.
-!- SIMILARITY: CONTAINS 6 LEUCINE-RICH REPEATS (LRR).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

EMBL; L31966; AAA86933.1; -.
EMBL; AF025377; AAC24981.1; -.
HSSP; P23945; 1XUN.
InterPro; IPR000276; GPCR_Rhodpsn.
InterPro; IPR001611; LRR.
InterPro; IPR000372; LRR_Nterm.
Pfam; PF00001; 7tm_1; 1.
Pfam; PF00560; LRR; 2.
Pfam; PF01462; LRRNT; 1.
SMART; SM00013; LRRNT; 1.
PROSITE; PS00237; G_PROTEIN_RECEP_Fl_1; 1.
PROSITE; PS50262; G_PROTEIN_RECEP_Fl_2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
Phosphorylation; Repeat; Leucine-rich repeat.
SIGNAL 1 17 POTENTIAL.
CHAIN 18 695 FOLLICLE STIMULATING HORMONE RECEPTOR.
DOMAIN 18 366 EXTRACELLULAR (POTENTIAL).
TRANSMEM 367 387 1 (POTENTIAL).
DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
TRANSMEM 399 421 2 (POTENTIAL).
DOMAIN 422 443 EXTRACELLULAR (POTENTIAL).
TRANSMEM 444 465 3 (POTENTIAL).
DOMAIN 466 485 CYTOPLASMIC (POTENTIAL).
TRANSMEM 486 508 4 (POTENTIAL).
DOMAIN 509 528 EXTRACELLULAR (POTENTIAL).
TRANSMEM 529 550 5 (POTENTIAL).
DOMAIN 551 573 CYTOPLASMIC (POTENTIAL).
TRANSMEM 574 597 6 (POTENTIAL).
DOMAIN 598 608 EXTRACELLULAR (POTENTIAL).
TRANSMEM 609 630 7 (POTENTIAL).
DOMAIN 631 695 CYTOPLASMIC (POTENTIAL).
REPEAT 44 68 LRR 1.
REPEAT 69 93 LRR 2.
REPEAT 119 143 LRR 3.
REPEAT 170 192 LRR 4.
REPEAT 193 216 LRR 5.
REPEAT 218 240 LRR 6.
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
S -> A (IN REF. 1).
T -> S (IN REF. 1).
V -> A (IN REF. 1).
V -> M (IN REF. 1).
Q -> H (IN REF. 1).
K -> R (IN REF. 1).
S -> T (IN REF. 1).
D -> N (IN REF. 1).
E -> K (IN REF. 1).
T -> A (IN REF. 1).
V -> E (IN REF. 1).
A -> T (IN REF. 1).
V -> I (IN REF. 1).

FT	CONFLICT	427	427	T -> S (IN REF. 1).
FT	CONFLICT	435	435	D -> N (IN REF. 1).
FT	CONFLICT	483	483	L -> V (IN REF. 1).
FT	CONFLICT	550	550	T -> I (IN REF. 1).
FT	CONFLICT	586	586	A -> V (IN REF. 1).
FT	CONFLICT	607	607	S -> L (IN REF. 1).
FT	CONFLICT	691	691	R -> H (IN REF. 1).
SQ	SEQUENCE	695 AA;	78172 MW;	E9EBEDB29C79C450 CRC64;

Query Match 4.4%; Score 104.5; DB 1; Length 695;
Best Local Similarity 16.5%; Pred. No. 4.1;
Matches 91; Conservative 91; Mismatches 169; Indels 199; Gaps 21;

QY	18	LPNG-INGIKDARKVTGVIGSGDFAKSLTIRLCRGYHVVGSRNPKFASEFFPHVVDV	76
Db	61	IPKGAFGGDLLEKI-----EISQNDVLEVEIAN-----VFSNLPKL-----	97
QY	77	THHEDALTKTNIIFVAIHREHYTSLWDLRHLLVG-----KILIDVSN 119	
Db	98	--HEIRIEKANNL-LYIDPDAFQNLPLNRYLLISNTGVKHLPAVHKIQSLQKVLDDIQDN	154
QY	120	MRINOYPESNAEYLASLPFDSLIVKGFNVSAWALQLGPKDASRQVYICSNNIOARQOVI	179
Db	155	INIH-----TVERNSFVGLSFESMILWL-----SKNGIREIHNCA-----	189
QY	180	ELARQLNFIPIDLGSLSSAREIENPLRLFTLWRGPVVVAISLATFFFLYSFVDVIHPY	239
Db	190	-----FNGTQLDELNSDNDNLEELPNDVFOGASGPVILDISRTHSLPSYLENKKL	244
QY	240	ARNQSDFYKIP-----IEIVNKT-----PIVAITLL-----	267
Db	245	RAKSTYNLKKLPLEKFTVTLMEASLTYPHSHCCAFANWRROIIDLHPICKNSILRQEVDM	304
QY	268	-----SLVYLAGLLAAAYQLYGYTKYRRF-----PPWLETWLQCRKQLG---	306
Db	305	TQARGQVSLAEDGESSLAKEFDTMYSFDYDLCNEVVDVICSPEDTFNPCEIDIMGHI	364
QY	307	--LLSFFFMVHVA-----YSLCLPMRRSERYLFLNMAYQQVHANIEWSNE	351
Db	365	LRVLWIFISILAITGNIIVLILITSQYKLTVP-----RFLMCNLAFAD-----	408
QY	352	EEVWRIEMYISFGIMSLGLLSLLAVTSIPSVSNALNWRE-----FSFIQSTLG---	399
Db	409	-----LCIGIYLLIASVDIHTKTQYHNYAIDWQTGAGCDAAGFTTFVASELSVYT	459
QY	400	YVALLISTFH-----VLIYGWKRAFEERYRYTPPNFV-LALVLP	439
Db	460	LTAITLERWHTITHAMQLQCKVQLRHAASIMLVGVFAFTVALFPFIIGISSYMKVSI	519
QY	440	SIVILDLLQL	449
Db	520	MDIDSPLSQL	529

RESULT 11
SOTB_ECO57
ID SOTB_ECO57 STANDARD; PRT; 396 AA.
AC P58529;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR Z2173 OR ECS2135.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,


```
RX MEDLINE=99369894; PubMed=10438792;
RA Carole S., Pichoff S., Bouche J.-P.;
RT "Escherichia coli gene ydeA encodes a major facilitator pump which
RT exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside.";
RL J. Bacteriol. 181:5123-5125(1999).
CC -!- FUNCTION: Involved in the efflux of sugars. The physiological role
CC may be the reduction of the intracellular concentration of toxic
CC sugars or sugar metabolites. Transports L-arabinose and to a
CC lesser extent IPTG. Seems to contribute to the control of the
CC arabinose regulon.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS
CC THE DRUG RESISTANCE TRANSLOCASE FAMILY). SOTB (TC 2.A.1.2)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000250; AAC74601.1; -.
DR EMBL; D90795; BAA15210.1; -.
DR EMBL; D90796; BAA15218.1; -.
DR EMBL; D90797; BAA15230.1; -.
DR EMBL; M96235; -; NOT_ANNOTATED_CDS.
DR EcoGene; Egl1636; sotB.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42538 MW; CB6A34CA4EE6D4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 2.3;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTGVIGSGDPAKSLTIRLCRGYHVVGSRNPKFASEFFPHVVDVTHHEDALTKTNIIF 90
Db 31 VPVGLL--SDIAQSFHMQTAQVGIMLTIYAWVVALMSLPFMLMTSQVERRKLLICLFVVF 88
QY 91 VAIHREHYTSLWDLRHLVGLKILIDVSNMNRINQYPESNAEYLASLPDLSLVKGFNVVS 150
Db 89 IASHVLSELS-WSFTVLVISRI-----GVAFAHAIF-----WSITA 123
QY 151 AWALQLGP--KDasRQVYICSNNIQARQVIELAR-----QLNFIPIDLGLS----- 195
Db 124 SLAIRMAPAGKRAQALSLIATGTALAMVLGLPLGRIVGQYFGWRMTFFAIGALITLTC 183
QY 196 -----SSAREIENLPLRFTLWRGPVVVAISLAT-----FFFLYSFVRDVIHPY 239
Db 184 LIKLLPLLPSEHSGSKSLPL----LFRPALMSIYLLTVVVVFAHYTAYS-----IEPF 235
QY 240 ARN---QOSDFYKIPiEIVNKTLPiVAITLLSLVYLAGLLAAAYQLYGTKYRRFPWWLE 296
Db 236 VQNIAGFSANF-----ATALLLLGGAGIIGSVIFGKLGNYAS----- 274
QY 297 TWLQCRKQLGLLSFFFAVMHVAYSICLPMRRSERYLFLNMAQQVQHANIENSWNEEEVVR 356
```

```
Db 275 -----ALVSTAIALLLVCLALLPAANSE-----IHLGVLSIF-----WG 309
QY 357 IEMYISFGIMSLGLLSL-----LAVTSIPSVSN-----ALNWREFSFIQST 397
Db 310 IAMMIIGLGMQVKVLALAPDATDVAMALFSGIFNIGIGAGALVGNQVSLHWS-----MSM 364
QY 398 LGYVALLISTFHVLIYG-----WKRAFEFE 422
Db 365 IGYVG-AVPAFAALIWSIIIFRRWPVTLEEQ 394

RESULT 13
NPT1_MOUSE
ID NPT1_MOUSE STANDARD; PRT; 465 AA.
AC Q61983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Renal sodium-dependent phosphate transport protein 1 (Sodium/phosphate
DE cotransporter 1) (Na(+)/PI cotransporter 1) (Renal sodium-phosphate
DE transport protein 1) (Renal Na(+)-dependent phosphate cotransporter
DE 1).
GN SLC17A1 OR NPT1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95335846; PubMed=7611445;
RA Chong S.S., Kozak C.A., Liu L., Kristjansson K., Dunn S.T.,
RA Bourdeau J.E., Hughes M.R.;
RT "Cloning, genetic mapping, and expression analysis of a mouse renal
RT sodium-dependent phosphate cotransporter.";
RL Am. J. Physiol. 268:F1038-F1045(1995).
CC -!- FUNCTION: IMPORTANT FOR THE RESORPTION OF PHOSPHATE BY THE KIDNEY.
CC MAY BE INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE INTO CELLS VIA
CC NA+ COTRANSPORT IN THE RENAL BRUSH BORDER MEMBRANE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: KIDNEY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77241; CAA54459.1; -.
DR MGD; MGI:103209; SLC17a1.
DR InterPro; IPR004745; Pi_cotransport.
DR TIGRFAMS; TIGR00894; 2A0114euk; 1.
KW Transport; Symport; Sodium transport; Transmembrane; Glycoprotein.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 337 356 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 465 AA; 51589 MW; C67EE25A2C291EEF CRC64;

Query Match 4.4%; Score 103; DB 1; Length 465;
Best Local Similarity 18.7%; Pred. No. 3.3;
```


Db	303	INNKLPIKINIMDIKNNYWLAGFTAADGSFSLSSMYNPKDTLFLFKDM-----	347
QY	162	SROVYICSNNIQARQOVIELARQLNFIPIDLGSLSSAREIENPLRLFTLWRGPPVVAIS	221
Db	348	-RPSYVIS-QVETRKELIYLIQE-----SFDL-SISNVKKVGNRKLKDFKLFTRTTDELMK	400
QY	222	LATEFFELYSFVRDVIHPYARNQOSDYPKIP----IEIVNKTLPIVAITLLSLVYLAGLLA	277
Db	401	-----FIYYF--DKFLPLHDNKOPNYIKFRFNTFIKSYNNWNNRVFGLVLS--YINNIKI	451
QY	278	AAAYQLYYGTTY	288
Db	452	DNYDYYYYNKY	462

Search completed: January 17, 2003, 19:01:20
Job time : 20 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 18:58:14 ; Search time 21 seconds
(without alignments)
2078.335 Million cell updates/sec

Title: US-09-455-486-6
Perfect score: 2351
Sequence: 1 MESISMGMSPKSLSETCLPN.....ALVLPISIVILDLLQLCRYPD 454

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160	6.8	198	D95285	conserved hypothet
2	159.5	6.8	211	AC2560	hypothetical prote
3	156.5	6.7	239	T50571	probable oxidoredu
4	147	6.3	213	H69400	conserved hypothet
5	142	6.0	212	D69361	conserved hypothet
6	142	6.0	224	T10120	F420-dependent NAD
7	133	5.7	223	D64487	hypothetical prote
8	121	5.1	191	F86826	hypothetical prote
9	120.5	5.1	232	A69131	conserved hypothet
10	119.5	5.1	216	T00121	hypothetical prote
11	116.5	5.0	242	G82642	conserved hypothet
12	116	4.9	198	AB3182	conserved hypothet
13	111.5	4.7	695	JN0898	folilitropin recept
14	111.5	4.7	1228	S59681	probable membrane
15	110.5	4.7	222	B84410	hypothetical prote
16	110.5	4.7	694	JC4301	folilitropin recept
17	109	4.6	320	T28379	ORF MSV218 hypothe
18	109	4.6	712	S50969	probable membrane
19	109	4.6	1184	H71436	hypothetical prote
20	109	4.6	1301	D85188	disease resistance
21	107.5	4.6	629	S60385	probable membrane
22	107.5	4.6	695	QRHUFT	folilitropin recept
23	105	4.5	442	B64582	sodium- and chlori
24	105	4.5	604	T31042	hypothetical prote
25	104	4.4	396	C64907	chloramphenicol re
26	104	4.4	396	G90895	hypothetical prote
27	104	4.4	396	H85721	probable resistanc
28	103	4.4	465	S69915	sodium-phosphate t
29	102.5	4.4	314	H98310	probable peptide A

30	102.5	4.4	314	2	AB2972	hypothetical prote
31	102.5	4.4	1242	2	T39453	probable mrna stab
32	102	4.3	220	2	AG3547	bicyclomycin resis
33	102	4.3	574	2	T41068	hypothetical prote
34	101	4.3	320	2	E71139	hypothetical prote
35	101	4.3	348	2	T12384	NADH2 dehydrogenas
36	101	4.3	420	2	F69144	O-antigen transpor
37	101	4.3	501	2	T02134	hypothetical prote
38	101	4.3	503	2	C86250	hypothetical prote
39	101	4.3	735	2	A83006	hypothetical prote
40	100.5	4.3	346	2	T11181	NADH2 dehydrogenas
41	100.5	4.3	1780	2	A85045	probable glucan sy
42	100	4.3	395	2	C71219	hypothetical prote
43	100	4.3	476	2	A28439	endonuclease Scel
44	100	4.3	714	2	AF2479	ABC transporter AT
45	99.5	4.2	452	2	C71391	NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

D95285
conserved hypothetical protein Sma0349 [imported] - Sinorhizobium meliloti (strain 10
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: D95285
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: D95285
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-198 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK64846.1; PID:g14523260; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSyma
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0349
A;Genome: plasmid

Query Match 6.8%; Score 160; DB 2; Length 198;
Best Local Similarity 26.7%; Pred. No. 2.1e-05;
Matches 55; Conservative 40; Mismatches 83; Indels 28; Gaps 8;

QY	32	TVGVIGSGDFAKSLTIRLIRCGYHVVI-GSRNPKFAS---EFFPHVVDVTHHEDALTKTN	87
Db	3	TYAIIAGAGIAGSALAERTFAAQIPAIIANSRGPASLSSTDRFGASVKAVELKDAL-QAD	61
QY	88	IIFVAIHREHYTSLWDLRLHLV---GKILIDVSNMNRINQYP-----ESNAEYLASLFP	138
Db	62	VVILAV---PYDSIADIVTQVSDWGGQIVVDASNAIDFAFKPRDLGGRLSTEIVSELVP	118
QY	139	DSLIVKGFNVVSAWALQLGPK--DASRQVYICSNNIQAQQVIELARQLNFIPIDLGSLS	196
Db	119	GAKVVKAFTNLPAAVLAADPKGTGSRVFLSGNHSNDANRQVAELISSLGFAFVDLGTLA	178
QY	197	SARETENPLRLFTLWRGPVVVAISL	222
Db	179	ASGPIQQF-----GRPLVALNL	195

RESULT 2

```

QY      132 YLASLFPDSLIVKGFNVVSAWALQ-----LGPKDASRQVYICSNNIQARQQV 178
           ||| ||| :   :||| |||               || : | : :
Db      143 QAAALLPDSRVAAAFHLSAVLLQDPDEIDETDVMVLGERADVEI-----VOA---- 192
                                           :||
QY      179 IELARQLNFIP-----IDLGSLSSAREIENPLRLFTLWR 213
           || : ||| : | | :| :| :| :| :| :|
Db      193 --LAGR---IPGMRGVFAGRLRNAHQVESLVANLISVNR 226
                               :| :| :| :| :| :| :| :|

RESULT 4
H69400
conserved hypothetical protein AF1209 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: H69400
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum
; Fleischmann, R.D.; Quackenbush

```

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,
A;Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: H69400
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A;Residues: 1-213 <KLE>
A;Cross-references: GB:AE001021; GB:AF000782; NID:g2689344; PIDN:AAB90038.1; PID:g264
C;Superfamily: conserved hypothetical protein MJ1501

Query Match 6.3%; Score 147; DB 2; Length 213;

Best Local Similarity 27.5%; Pred. NO. 0.00024;;

Matches 56; Conservative 44; Mismatches 68; Indels 36; Gaps 11;

oy 37 GSGDFAKSLTIRLCYGHVVIGSRN-----PKFASFFPHWVDVT---HHEDALTKTNI 88

b 8 GTGNLGEGLARWGKLGYEIVGSRKLEKAELADSYLKVKVCDASTICGVWDATITETTTT---

```

b 67 --VAV-----FTIPWEFAFDTAEMLKRLQLAGVKVISPLVPMKKVDNF-VYVRPEGSAAE 119
Y 132 YLASLFPDSLIVKGFNVVSAWALQLGPKDASRQVYICSNNIQARQVIELARQLNFI-PI 190
b 120 KLASVLEESSVVAAYHSIPARRFANLGEFEWDVPICGDS-GAKEVVVDLTEKISGLRAL 178
Y 191 DLGSLSSAREIENL-PLRLFLLWR 213
c 179 DAGGLSNAHLVESLTPLIINVMKR 202

RESULT 5
9361
unserved hypothetical protein AF0892 - Archaeoglobus fulgidus
Species: Archaeoglobus fulgidus
Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
Accession: D69361
Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.;
Fleischmann, R.D.

```

Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, L.H.O.; Woese, C.R.; Venter, J.C.
 Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaean, *Ignicoccus hospitalis*.
 Reference number: A69250; MUID:98049343; PMID:9389475
 Accession: D69361
 Status: preliminary; nucleic acid sequence not shown; translation not shown
 Molecule type: DNA
 Residues: 1-212 <LE>
 Cross-references: GB:AE001042; GB:AE000782; NID:g2689365; PIDN:AAB90348.1; PID:g264
 Superfamily: conserved hypothetical protein

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Delta H)
C/Species: Methanobacterium thermoautotrophicum
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C/Accession: A69131
R/Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A/Reference number: A69000; MUID:98037514; PMID:9371463
A/Accession: A69131
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-232 <MTH>
A/Cross-references: GB:AE000811; GB:AE000666; NID:g2621287; PIDN:AAB84754.1; PID:g262129
A/Experimental source: strain Delta H
C/Genetics:
A/Gene: MTH248
A/Start codon: TTG
C/Superfamily: conserved hypothetical protein MJ1501

Query Match	5.1%	Score 120.5;	DB 2;	Length 232;
Best Local Similarity	23.8%;	Pred. No. 0.034;		
Matches 49;	Conservative 42;	Mismatches 90;	Indels 25;	Gaps 6;

QY 30 KVTVGISGDEFAKSLTIRLRCGYHVIVIGSRNPKFASEFFPHVDVTHHE----- 80

Db 8 KIAV-IGGTGDOGGLGLALRFAVAGEEVIIGSRDAEKASKAASKVLEIAGRDDISVEGATN 66

QY 81 -DALTKTNIIFVAIH-REHYTSLWDLRHLLVGKILIDV-----NNMRINQYPESN 129

Db 67 PDAAASADVWVLTPLQAQNVTLASIRDQVRDKVLIDATVPIDSCIGGSVAVRYIDLWEGS 126

Qy 130 AEYLALF---PDSLIVGFNVSAWALQLGPKDASQVYICSNNIQARQVIELARQLN 186

```

Db 127 AAERARFLREQGTRVAAAFNNISASALLEVSEVPDCLVSDHRDALEVAFLAEKID 186

```

QY 187 FI-PIDGLSSAREIENPLRLFTL 211

Db 187 GVRAIECGGLENARIIEKITPLINL 212

RESULT 10
T00121
hypothetical protein 8 - Leptospira interrogans
C:Species: Leptospira interrogans
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00121
R:Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.
Gene 215, 37-45, 1998
A:Title: Physical and genetic maps of the Leptospira interrogans serovar ic

A;Reference number: Z14115; MUID:98332717; PMID:9666070
 A;Accession: T00121
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-216 <TAK>
 A;Cross-references: EMBL:AB010203; NID:g2780763; PIDN:BAA24368.1; PID:g2780765
 A;Experimental source: strain Ictero No.1; substrain icterohaemorrhagiae

Query Match	5.1%;	Score 119.5;	DB 2;	Length 216;
Best Local Similarity	23.7%;	Pred. No. 0.037;		
Matches	47;	Conservative	37;	Mismatches 91;
				Indels 23;
				Gaps 6;

33 VGVIGSGDEAKSLTIRLCRGYHVVIGSRNPKFASEFFPHV ---VDVTHHEDALTKTNII 89

6 IGILGSGIVGQTLANGFLKYGAEVKYIGTRDFGKLDWLAKAGAGASIGSFSEANFGEII 65

90 FVAIHREHYTSLWDLRHL--LVGKILIDVSNMR-----INQYPESNAEYLASL 136

db 66 VLCSKGSVASEVLTSLGSDLSNGKTIIDTNTPTSEIPPQNGVLNFTTSYNSLMEKLRQK 125

137 FPD SLVKGENVVS AWALQLGP--KDSRQVYVICSNNIQRQQVIELRQLNFIPIDLGS 194

Db 126 APKANFVKCFSSVGS-GLMVNPQLKGEKPSMFICGNDSDSSKKQIKEILDTFGWDTEDMGK 184

QY 195 LSSAREIENPLRLFTLW 212
: : | | | : | |

Db 185 VEAARAIEPLCI----LW 198

RESULT 11
G82642

conserved hypothetical protein XF1737 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: G82642
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide S
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below

Query Match	5.0%;	Score 116.5;	DB 2;	Length 242;
Best Local Similarity	24.0%;	Pred. No. 0.075;		
Matches 46;	Conservative	40;	Mismatches 85;	Indels 21;
				Gaps 7;

QY 28 ARKVTGVIGSGDFAKSLTIRLIRCGYHVIGSRNP---KFASEFFPHVVDVTHEDAL 83

Dbb 30 AAPMRIGVIGAGSLGGTVGRLWVKACHEVMFSSRNPDKLEAMARELEPR-ASVGQPLAAT 88

84 TKTNIIFVAIHREHYTSLW-DLRHLLVGKILIDVSNMNRINQ---YPESN-----AEYLAS 135

89 EFGTVLLAVPFEPALQVGRDLRSAYRGKIVLDSINPWGASSADVYREARELGVQATVVK 148

136 LFPDSLIVKGFNVVSAWALQLGPKDASRQ-----VYICSNNIARQQVIELARQLNFIPI 190

Dbb 149 YMPGARLVRAFSVDATWE--TSASRRGGRRIGNPLASDDAEAMKVAEGLVRDAGCDPV 205

QY 191 DLGSLSSAREIE 202
:|:|:|:

db 206 IVGNLAAASFQ 217

RESULT 12
AB3182

```
conserved hypothetical protein Atu5183 [imported] - Agrobacterium tumefaciens (strain
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002_
C;Accession: AB3182
```

R.; Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Wood-
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A; Genome: plasmid

Db 1 MTVGII GAGNIGGAFATALGKAGTEAVTANSRGPESTAIIVSKI.GSTTRAGSVPEAAAAOA 60

61 -IVIVAVP-----WSKTPGALAGLNEGDBRIVTDANNSTEAPIYRPADI.GGRTSTDTET 112

Db 113 ALVPGARVKAEN-----HLTPKQLSGDPHSEGGRRVLFYSGDDMRKAKEVGAIIDRI 165

Db 166 GFFGIDLGLPVGSQLQFP-----GGPLPALNLVKF 197

C;Accession: JN0898; S36452

A;Reference number: JN0898; MUID:94071854; PMID:7504463

A;Note: the authors translated the codon AGT for residue 488 as Arg

F;1-17/Domain: signal sequence #status predicted <SIG>

F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR

F:191.199.293.318/Binding site: carbohydrate (Asp) (covalent) #status predicted

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

F:720-736

;720-736/Domain: transmembrane #status predicted <TM3>

Query Match									
Best Local Similarity 4.78; Score 111.5; DB 2; Length 1228;									
Matches 111; Conservative 71; Mismatches 152; Indels 225; Gaps 28;									
QY	11	KSLSETCLPNGINGIKDARK-----	VTGVIGSGDFAKSLTIRLI	50					
Db	292	KTMAETTISSGLAENKYLRLVDTIFALKPSNVDTLLTKSWIAVVIKGMSTYATHOPLKAL	351						
QY	51	R-----CGYHVVIGSRNPKF---ASEFFPHVVDVTHHEDAL-----	83						
Db	352	RKIPGVFHIMCTY---LASETPEVYQAASQCLISILSESVKDDLLLYTPSDEKVFKNVD	408						
QY	84	-----TKTNIIFVAIHREHYTSLWDLRHLLVGKILIDVSNMNRINQYPESNAEYLASLF	137						
Db	409	EIISQIAKTFIDFLSIRYSHCS-----REIL--KILVAAFNKFRY-----RSNPHFLKSL-	456						
QY	138	PDSLIVKGFNVSAWALQGLPKDASRQVYICSNNIQARQQVIELARQLNFI-----	188						
Db	457	-----KIVDTWRVN-----EQFMDLRNEIELVIGASISAMG	488						
QY	189	-----PIDGLSSLSSAREIENLPLRLFTLWRGPV-----VAISLATEFFFLYSFVRDVI	236						
Db	489	PEMILAEAPLNDNPSSER-----PGR---AWLLPLIRDYTKNANLATE-----QNEL	533						
QY	237	HPYARNQOSDFYKIPIEIVN-----KTLP-----IVAITLLSL	269						
Db	534	APIYKSFQSKFDKVPESIQLRVFQTIQVQIWSLTPRFCELPMDLRESFDEFASELSSL	593						
QY	270	VYL-----AGLLAAAYQLYYGTK-----YRRFP-----PWLETWLQCRK	303						
Db	594	LYSEVELRTTICHALKVLAESNVSYAESSHNVLLQLQRFPISEAQNIEYLST-----K	648						
QY	304	QLGLLSFFFAMVHVAYSCLCPMRRS-----ERYL-----FLNMAYQQVHANIEWSNE	351						
Db	649	STNLLAVLFNV---YTQTPNARSYILETIDQYLKITSKEDLEKTFENNVCGLLKNSMNE	704						
QY	352	EEVWRIEMVISFGIMSLGLLSLL--AVTSIPSVSNALNWRFSFIQSTGLGYVALLISTFH	409						
Db	705	ESSGNVNKEKKPKQLTATLLDLIICMITYLPVSSYSALFSMFS-----LTVNSAD	754						
QY	410	VLIYGWKRAFEYYRYFT	428						
Db	755	ALIQ--KRA-----YRIIT	766						
RESULT 15									
B84410									
hypothetical protein vng2607c [imported] - Halobacterium sp. NRC-1									
C;Species: Halobacterium sp. NRC-1									
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001									
C;Accession: B84410									
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S									
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl									
Jung, K.H.; Alam, M.; Freitas, T.									
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181; 2000									
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li									
A;Title: Genome sequence of Halobacterium species NRC-1.									
A;Reference number: A84160; MUID:20504483; PMID:11016950									
A;Accession: B84410									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-222 <STC>									
A;Cross-references: GB:AE004437; NID:g10581987; PIDN:AAG20646.1; GSPDB:GN00138									
C;Genetics:									
A;Gene: VNG2607C									
C;Superfamily: conserved hypothetical protein MJ1501									
Query Match									
Best Local Similarity 4.78; Score 110.5; DB 2; Length 222;									
Matches 49; Conservative 34; Mismatches 94; Indels 25; Gaps 6;									
QY	37	GSGDFAKSLTIRLCGYH-VVIGSRNPKFASEFFPHVVDVT-----HHEDAL	83						

Db	8	GTGDIGAGLALRWATSDHDDIVIGSRDPEKARETAAAYEDTLADQGVDRKLTGFANEMAA	67
QY	84	TKTNIIFVAIHREHYTSLW-----DLRHLLVGKILIDVSNMNRINQYPESNAE---YL	133
Db	68	DRADVVVAVPAYHVTDVMGAVADRLDADTLVISPAVGIASGEHGLHYNPPSAGSVTALV	127
QY	134	ASLEPDSL-IVKGFNVSAWALQGLPKDASRQVYICSNNIQARQQVIELARQLNFI-PID	191
Db	128	ADAAPDGVVVGAFHNLAADRLADLTDELADATLVVGNDEGARTVAELADDDITGLRALD	187
QY	192	LGSLSAREIENLPLRLFTLWR	213
Db	188	AGPVENAAEVESLTPLLINLAR	209

Search completed: January 17, 2003, 19:00:16
Job time : 33 secs